

# ORTHONORMAL BASES OF REGULAR WAVELETS IN SPACES OF HOMOGENEOUS TYPE

PASCAL AUSCHER AND TUOMAS HYTÖNEN

**ABSTRACT.** Adapting the recently developed randomized dyadic structures, we introduce the notion of spline function in geometrically doubling quasi-metric spaces. Such functions have interpolation and reproducing properties as the linear splines in Euclidean spaces. They also have Hölder regularity. This is used to build an orthonormal basis of Hölder-continuous wavelets with exponential decay in any space of homogeneous type. As in the classical theory, wavelet bases provide a universal Calderón reproducing formula to study and develop function space theory and singular integrals. We discuss the examples of  $L^p$  spaces, BMO and apply this to a proof of the  $T(1)$  theorem. As no extra condition (like ‘reverse doubling’, ‘small boundary’ of balls, etc.) on the space of homogeneous type is required, our results extend a long line of works on the subject.

## 1. INTRODUCTION

The main goals of this paper are two-fold: the construction of orthonormal, Hölder-continuous wavelet bases in general spaces of homogeneous type, and their applications in the theory of singular integrals and function spaces in the same general set-up. Despite several existing results of related nature, the scope of our theory is completely new in at least two respects:

First, as far as we are aware, it seems that we offer the first construction of an orthonormal wavelet *basis*, as opposed to a *frame*, in this setting. Only relatively recently, Deng and Han wrote [11, p. 40]: “Orthonormal wavelet bases are out of reach on a space of homogeneous type. Instead the theory of frames will be used. Roughly speaking using a frame means that you tolerate a limited amount of redundancy while redundancy is completely avoided with a basis.” Here, we build a genuine basis, and so avoid this redundancy. Our construction starts from appropriate *splines*, which also seem to be new on an abstract quasi-metric, even metric, space, and of independent interest.

Second, we are careful not to impose any additional assumptions other than those defining a space of homogeneous type in the sense of Coifman and Weiss [6]—that is, a set  $X$  equipped with a quasi-distance  $d$  and a Borel measure  $\mu$  that is doubling on the quasi-metric balls—, and we keep working with the original given  $d$ , without changing to an ‘equivalent’ one (see below). The difficulty is that the quasi-distance, in contrast to a distance, may not be Hölder-regular, and quasi-metric balls might not be open, nor even Borel sets with respect to the topology defined by the quasi-distance (openness is assumed in [6]). We merely assume that balls be Borel sets, but even this assumption can be easily remedied by. See Section 5.

In contrast to this, we emphasize that a substantial part of the rich literature on analysis in spaces of homogeneous type involves additional assumptions on the underlying space. Let us consider in particular the existence of a Littlewood–Paley decomposition (or a Calderón reproducing formula) with a *regular* kernel, which has been the object of numerous works. (Decompositions with discontinuous kernels arising from martingale differences are well-known even in more general set-ups; this is a different story.) Indeed, the regular Littlewood–Paley decomposition is the core formula that allows the various characterizations of Hardy spaces, the development of function spaces of Besov and Triebel–Lizorkin type, and the analysis of operators that act on them. (Here, we have in mind a definition of Besov spaces linked with the modulus of continuity and thereby

---

2010 *Mathematics Subject Classification.* 42C40, 41A15, 30Lxx, 42B25.

*Key words and phrases.* Geometrically doubling quasi-metric space, space of homogeneous type, spline function, wavelet, orthonormal basis, dyadic cube, random geometric construction,  $T(1)$  theorem.

the (quasi-)metric of the underlying space; some other generalizations from the point-of-view of abstract approximation spaces, as in [26], are yet another story.)

As far as we understand, there has always been some kind of further hypothesis either on  $X$ ,  $d$  or  $\mu$ . The seminal work in this context is the paper of David–Journé–Semmes on the  $T(b)$  theorem [9]. The starting point of their theory is the existence of a new Hölder-continuous quasi-metric adapted to the measure  $\mu$  and *topologically* (but not geometrically) equivalent to the original  $d$ , which is provided by work of Macías–Segovia [25, Theorems 2 and 3] (for more precise versions of the Macías–Segovia results, see [34] and [31]). Their singular integrals are then defined relative to this new quasi-metric, which leads to a possibly different class of Calderón–Zygmund operators than those defined in terms of the original  $d$  and  $\mu$ . (Even then, [9] still needs some more technical assumptions such as no point masses, infinite mass for  $X$  and even a “small boundary” property for balls.) Working with this kind of new quasi-metric has become a common set-up in the literature. The theory of function spaces under this assumption has been developed by Han–Sawyer [16], see also the nice review by Han–Weiss [17].

By [25], Theorem 2 (for an elementary proof, see [34]), it is possible to only change the quasi-metric to a metrically equivalent one that is Hölder-regular. This does not change the classes of singular integral operators (up to changing the Hölder-exponents), but leaves another difficulty: some estimates require a control from below for the growth of balls. See the discussion below about (1.1). In several more recent papers [14, 15, 42, 43] on function spaces on (quasi-)metric spaces, this issue is circumvented by assuming the ‘reverse doubling’ property

$$\mu(B(x, Cr)) \geq (1 + \varepsilon)\mu(B(x, r)), \quad 0 < r < \text{diam}(X)/C,$$

which is equivalent to the non-emptiness of annuli:  $B(x, R) \setminus B(x, r) \neq \emptyset$  whenever  $R/r$  is large enough and  $r$  is at most a fraction of  $\text{diam}(X)$ . This latter condition is another common assumption; it was only recently eliminated from some results about positive integral operators on metric spaces by Kairema [22].

The reverse doubling excludes in particular the presence of point masses (which is also frequently assumed, even if empty annuli are otherwise tolerated), and therefore rules out some basic examples like discrete groups  $\mathbb{Z}$  or  $\mathbb{Z}/p\mathbb{Z}$  and the multidimensional analogs, the typical discrete metric structures arising in theoretical computer science (trees, graphs, or strings from a finite alphabet),  $\mathbb{Q}_p$  from arithmetics, or discrete approximations of other spaces of homogeneous type (like those constructed in [2]), even if the original space did satisfy the reverse doubling property and some *ad hoc* orthonormal (or bi-orthogonal) wavelet bases had been constructed. To mention a few references in these directions, see [3], [35], [37]. We remark that even if Hölder-regularity on discrete structures might not be an issue at first glance, it becomes one thinking the discrete structure as approximating a continuous one.

In contrast, we repeat, we do not make any such additional assumptions or changes to a different quasi-metric.

Let us come to the construction of spline wavelets. Splines have a long history in approximation theory, due in particular to the interpolation property and their polynomial nature, and were popularized by early books in the 1960’s. (The first book listed from a MathSciNet title search “spline” is [1] and the reviewer dates the appearance of the splines to Schoenberg’s work during World War II; see [36].) As for spline wavelets, they were first constructed on the real line and on the 1-torus by Strömberg [38]. They were rediscovered independently by Battle [4] and Lemarié [23] in Euclidean spaces. Using the Multiresolution Analysis scheme of Mallat [27] and Meyer [28], it became clear that the existence of regular, compactly supported splines with the interpolation property and reproducing formula leads to bases of regular spline wavelets. This was attempted in other geometrical contexts. For example, splines and spline wavelets were constructed in [24] on a stratified Lie group following a minimisation procedure using the existence of a sublaplacian, and in [7] on compact manifolds like spheres. We do not claim to be exhaustive in this history. Nevertheless, we have to follow a new route as there is no group structure attached, nor any local coordinates to help. We rather use a probabilistic approach as described next.

Central to our analysis are the dyadic structures in a geometrically doubling quasi-metric space as constructed by Christ [5] (a construction in Ahlfors–David spaces was done earlier by David

[8]), and their randomized versions recently studied by the second author with Martikainen [21] and Kairema [19]. (Another variant is due to Nazarov, Reznikov and Volberg [32].) We build the splines as averages of the indicators of dyadic cubes under a random choice of the dyadic system. This can be motivated by an appropriate point of view at the classical piecewise linear splines on  $\mathbb{R}$ , which are generated by the function

$$s(x) = x1_{(0,1]}(x) + (2-x)1_{(1,2)}(x).$$

Observe that random dyadic intervals of sidelength 1, in the sense of Nazarov, Treil and Volberg [33, Sec. 9.1], can be defined by translating the standard intervals  $[k, k+1)$  by a random number  $u \in [0, 1)$ . Thus the unit cube with left end at the origin,  $[0, 1)$ , is translated to  $[u, u+1)$ . It is easy to check that the average of its indicator function  $1_{[u, u+1)}(x) = 1_{[0,1)}(x-u)$  over the uniform random choice of  $u$  is precisely

$$\int_0^1 1_{[0,1)}(x-u) du = 1_{[0,1)} * 1_{[0,1)}(x) = s(x).$$

So one can indeed think of splines as averages of the indicators of random dyadic intervals. This is the basic idea which will guide us in constructing splines for a quasi-metric space  $X$ .

Once the splines are available, and only at this point we add into our considerations the doubling Borel measure, we can obtain the orthonormal bases of scaling functions and the wavelets by the general procedure of Meyer. These will be Hölder-continuous functions  $\psi_\alpha^k$  with an *exponential localization* in the  $\delta^k$ -neighbourhood of a given point  $y_\alpha^k$ :

$$|\psi_\alpha^k(x)| \leq \frac{C}{\sqrt{\mu(B(y_\alpha^k, \delta^k))}} \exp\left(-\gamma\left(\frac{d(x, y_\alpha^k)}{\delta^k}\right)^a\right)$$

for some  $a \in (0, 1]$  that depends only on the quasi-triangle constant (with  $a = 1$  in the metric case or when the quasi-distance is Lipschitz-continuous). The construction of *boundedly supported* wavelets in this generality remains an open problem.

There is an important feature related to the collections of points  $\mathcal{Y}^k := \{y_\alpha^k\}$ , which index the wavelets of the scale  $\delta^k$ . These collections arise as  $\mathcal{Y}^k = \mathcal{X}^{k+1} \setminus \mathcal{X}^k$ , where  $\dots \subseteq \mathcal{X}^k \subseteq \mathcal{X}^{k+1} \subseteq \dots$  is an increasing sequence of point-sets, and each  $\mathcal{X}^k$  is both  $\delta^k$ -separated and (up to constant)  $\delta^k$ -dense in  $X$ . The important feature is that  $\mathcal{Y}^k$  can be much sparser than  $\mathcal{X}^{k+1}$  in some regions of the space, which reflects the ‘absence of the length scale  $\delta^{k+1}$ ’ in the local geometry of the space. Accordingly, the distance  $d(x, \mathcal{Y}^k)$  of a given element  $x \in X$  to the nearest point  $y_\alpha^k \in \mathcal{Y}^k$  will be a significant quantity, which can in general be much larger than  $\delta^k$ .

This quantity appears in central technical estimates when bounding series of the following type, which naturally arise in the context of function spaces:

$$(1.1) \quad \sum_{j \in \mathbb{Z}: \delta^j \geq r} \frac{1}{\mu(B(x, \delta^j))} \exp\left(-\gamma\left(\frac{d(x, \mathcal{Y}^j)}{\delta^j}\right)^a\right) \leq \frac{C}{\mu(B(x, r))}.$$

This (non-obvious but valid) estimate serves as a replacement of the following bound, which is repeatedly applied by Han, Müller and Yang in the reverse doubling context (cf. [15, Lemma 3.5]):

$$\sum_{j \in \mathbb{Z}: \delta^j \geq r} \frac{1}{\mu(B(x, \delta^j))} \leq \frac{C}{\mu(B(x, r))}.$$

Indeed, this is a quick consequence of reverse doubling, but invalid in general spaces of homogeneous type. A typical cause to destroy this latter estimate is the presence of a large empty annulus around  $x$ , but this then results in some large values of  $d(x, \mathcal{Y}^j)$ , which compensates for the failure of  $\mu(B(x, \delta^j))$  to grow fast enough in (1.1).

Although we do not develop the theory of function spaces in the detail of [15] here, it is clear that our analysis involving the quantity  $d(x, \mathcal{Y}^k)$  will extend to further related questions that we do not explicitly deal with; we believe that it could be used to extend large parts of the recent theory of ‘RD (reverse doubling) spaces’ to general spaces of homogeneous type. As an illustration, we give a short treatment the  $T(1)$  theorem using our construction.

Over the last decade or so, it has been discovered that several aspects of harmonic analysis can even be pushed beyond the setting of doubling measures; see e.g. [33, 40, 41] for some of the pioneering developments in the context of  $\mathbb{R}^d$ , and [20, 21] for recent developments in abstract quasi-metric spaces. We do not address the non-doubling measures here, as it seems that our regular splines and wavelets are most useful in the (already quite general) doubling situation. In fact, while our splines are constructed completely independently of any underlying measure, it turns out that they automatically have a good  $L^2$  theory with respect to any doubling measure on the space. This would not be the case for a non-doubling measure, and a successful wavelet theory for such measures, if possible at all, should probably be adapted to the particular measure in a more complicated manner, already at the early point, where we can manage with a purely geometric construction.

**About notation.** We use  $C$  to denote positive constants, whose value may change from one occurrence to the next. We also abbreviate ' $\leq C \times \dots$ ' to just ' $\lesssim$ '. We use  $\gamma$  in a similar role as a positive exponent. In contrast to  $C$ , it typically decreases from one occurrence to the next. For the measure ('volume') of balls  $B(x, r) := \{y \in X; d(x, y) < r\}$ , we sometimes use the abbreviations

$$V(x, r) := \mu(B(x, r)), \quad V(x, y) := V(x, d(x, y)).$$

**Acknowledgements.** This work started during a visit of the second author to Paris-Sud whose support is gratefully acknowledged. The second author was also partially supported by the Academy of Finland, grants 130166, 133264 and 218148. We thank Detlef Müller for answering our questions on his work on function spaces on Reverse Doubling spaces, and the anonymous referee for comments to improve the presentation.

## 2. PRELIMINARIES ON DYADIC SYSTEMS

**Introduction and motivation.** In this section we review the relevant-for-us parts of M. Christ's [5] construction of dyadic cubes in a space of homogeneous type, as well as its recent probabilistic version from [19, 21]. For the convenience of the reader, and since we feel that we have managed to slightly streamline the earlier presentations, we will give a self-contained treatment, even though the actual novelty is only in certain details of the construction.

Christ's cubes  $Q_\alpha^k$  (where  $k$  designates the generation or length scale of the cube, and  $\alpha$  indexes the cubes inside a given generation) are defined by two sets of auxiliary objects: the centre points  $x_\alpha^k$ , which determine the rough position of individual cubes, and a partial order  $\leq$  on the family of the index pairs  $(k, \alpha)$ , which determines the set inclusions among different cubes.

In [19, 21], a systematic method of constructing several Christ-type families of dyadic cubes  $Q_\alpha^k(\omega)$  was introduced, where  $\omega$  belongs to a parameter space  $\Omega$ . Here  $\Omega$  can be equipped with a probability measure, which gives rise to a notion of random dyadic cubes. This allows us to compute averages over a random choice of the cubes, and make probabilistic statements about them. As usual, averaging has a smoothing effect, and the random cubes will 'on average' enjoy better regularity properties than any fixed cubes would do. This is essential for our construction of Hölder-continuous splines.

Since the cubes  $Q_\alpha^k$  are determined by the dyadic points  $x_\alpha^k$  and the partial order  $\leq$ , the construction of the parametrized family of  $Q_\alpha^k(\omega)$  amounts to defining appropriate parameter-dependent points  $z_\alpha^k(\omega)$  and a parameter-dependent partial order  $\leq_\omega$ . This is achieved by first fixing the reference objects  $x_\alpha^k$  and  $\leq$  as in Christ's original work, and constructing  $z_\alpha^k(\omega)$  and  $\leq_\omega$  as their parametrized perturbations.

All these constructions involve certain arbitrary choices, which all work equally well for the applications considered in the earlier papers. A feature of our presentation here is that we are going to insist on somewhat more specific, rather than arbitrary, choices in certain details of the construction, as this will be a convenience later when using the dyadic cubes in the building of our splines.

We now turn to the details.

**General assumptions.** In what follows,  $X$  is a set equipped with a quasi-distance with quasi-triangle constant  $A_0 \geq 1$ , namely,  $d(x, y) = d(y, x) \geq 0$ ,  $d(x, y) = 0$  if and only if  $x = y$ , and  $d$  satisfies the quasi-triangle inequality

$$d(x, y) \leq A_0(d(x, z) + d(z, y)).$$

We assume that  $X$  has *the geometric doubling property*, namely that there exists a natural number  $N$  such that any given ball contains no more than  $N$  points at quasi-distance exceeding half its radius. This is a certain finite-dimensionality requirement on the space: for example, the infinite-dimensional Hilbert space  $\ell^2$  fails this property, since all the half-unit vectors  $\frac{1}{2}e_k$ ,  $k \in \mathbb{N}$ , belong to the unit-ball, while their mutual distance is  $\sqrt{2}/2 > 1/2$ . We do not need the doubling measure, or in fact any underlying measure, at this stage.

Let us fix a small parameter  $\delta > 0$ . For example, it suffices to take  $\delta \leq \frac{1}{1000}A_0^{-10}$ . Roughly speaking, the point is that phenomena on the length scale  $\delta^{k+1}$  should remain much smaller than the length scale  $\delta^k$ , even after repeated use of the quasi-triangle inequality where we ‘lose’ the constant  $A_0$  at every application. For example, if the points  $x_0, x_1, x_2 \dots$  satisfy  $d(x_{i-1}, x_i) < \delta^{k+1}$ , we can still conclude that  $d(x_0, x_r)$  is much less than  $\delta^k$  for  $r \leq 10$ .

**Reference dyadic points.** For every  $k \in \mathbb{Z}$ , we choose a set of *reference dyadic points*  $x_\alpha^k$  as follows: For  $k = 0$ , let  $\mathcal{X}^0 := \{x_\alpha^0\}_\alpha$  be a maximal collection of 1-separated points. Inductively, for  $k \in \mathbb{Z}_+$ , let  $\mathcal{X}^k := \{x_\alpha^k\}_\alpha \supseteq \mathcal{X}^{k-1}$  and  $\mathcal{X}^{-k} := \{x_\alpha^{-k}\}_\alpha \subseteq \mathcal{X}^{-(k-1)}$  be maximal  $\delta^k$ - and  $\delta^{-k}$ -separated collections in  $X$  and in  $\mathcal{X}^{-(k-1)}$ , respectively.

**2.1. Lemma.** *For all  $k \in \mathbb{Z}$  and  $x \in X$ , the reference dyadic points satisfy*

$$d(x_\alpha^k, x_\beta^k) \geq \delta^k \quad (\alpha \neq \beta), \quad d(x, \mathcal{X}^k) = \min_\alpha d(x, x_\alpha^k) < 2A_0\delta^k.$$

*Proof.* The separation property is part of the construction. By maximality, it follows that for all  $x \in X$  and  $k \geq 0$ ,

$$d(x, \mathcal{X}^k) = \min_\alpha d(x, x_\alpha^k) < \delta^k.$$

Also, given  $x \in X$ , we can recursively find points  $x_{\alpha_0}^0, x_{\alpha_1}^{-1}, \dots, x_{\alpha_k}^{-k}$  such that

$$d(x, x_{\alpha_0}^0) < 1, \quad d(x_{\alpha_0}^0, x_{\alpha_1}^{-1}) < \delta^{-1}, \quad \dots \quad d(x_{\alpha_{k-1}}^{-(k-1)}, x_{\alpha_k}^{-k}) < \delta^{-k},$$

and hence

$$d(x, x_{\alpha_k}^{-k}) < \sum_{j=0}^k A_0^{j+1} \delta^{-(k-j)} < \frac{A_0 \delta^{-k}}{1 - A_0 \delta} \leq 2A_0 \delta^{-k}. \quad \square$$

Note that  $\mathcal{X}^k \subseteq \mathcal{X}^{k+1}$ , so that every  $x_\alpha^k$  is also a point of the form  $x_\beta^{k+1}$ , and thus of all the finer levels. We denote  $\mathcal{Y}^k := \mathcal{X}^{k+1} \setminus \mathcal{X}^k$ , and relabel these points as  $\mathcal{Y}^k = \{y_\alpha^k\}_\alpha$ . These points will play an important role as a parameter set of our wavelets, to be constructed.

**Reference partial order.** We set up a *partial order*  $\leq$  among the pairs  $(k, \alpha)$  as follows: Each  $(k+1, \beta)$  satisfies  $(k+1, \beta) \leq (k, \alpha)$  for exactly one  $(k, \alpha)$ , in such a way that

$$(2.2) \quad d(x_\beta^{k+1}, x_\alpha^k) < \frac{1}{2A_0} \delta^k \implies (k+1, \beta) \leq (k, \alpha) \implies d(x_\beta^{k+1}, x_\alpha^k) < 2A_0 \delta^k.$$

The pairs  $(k+1, \beta)$  with  $(k+1, \beta) \leq (k, \alpha)$  are called the children of  $(k, \alpha)$ . Geometric doubling implies that their number is uniformly bounded. So far, we have essentially followed the original construction of M. Christ [5], a slight nuance being the choice of the point-sets  $\mathcal{X}^k$  in such a way as to have the nestedness  $\mathcal{X}^k \subseteq \mathcal{X}^{k+1}$ , which was not required in [5, 19, 21].

**Labels for the points.** For a successful perturbation argument to construct the parametrized dyadic points  $z_\alpha^k(\omega)$  below, we need certain book-keeping among the near-by dyadic points  $x_\alpha^k$  of the same generation.

Points  $(k, \alpha)$  and  $(k, \beta)$  are called *neighbours*, if they have children  $(k+1, \gamma) \leq (k, \alpha)$  and  $(k+1, \eta) \leq (k, \beta)$  such that  $d(x_\gamma^{k+1}, x_\eta^{k+1}) < (2A_0)^{-1}\delta^k$ . In this case

$$\begin{aligned} d(x_\alpha^k, x_\beta^k) &< A_0 d(x_\alpha^k, x_\gamma^{k+1}) + A_0^2 d(x_\gamma^{k+1}, x_\eta^{k+1}) + A_0^2 d(x_\eta^{k+1}, x_\beta^k) \\ &< 2A_0^2 \delta^k + \frac{1}{2}A_0 \delta^k + 2A_0^3 \delta^k < 5A_0^3 \delta^k. \end{aligned}$$

The number of neighbours that any point can have is also uniformly bounded.

We equip each pair  $(k, \alpha)$  with two *labels*, which are chosen from a finite set but which still locally distinguish between different pairs  $(k, \alpha)$ . The primary label, denoted by  $\text{label}_1(k, \alpha) \in \{0, 1, \dots, L\}$ , where  $L$  is the maximal number of neighbours, is chosen in such a way that any two neighbours have a different label. The secondary label, denoted by  $\text{label}_2(k, \alpha) \in \{1, \dots, M\}$ , where  $M$  is the maximal number of children, is chosen in such a way that no two children of the same parent have the same label. These labels were introduced in [19] with slightly different notation.

**Parametrized points and partial order.** As described above, we now want to perform a perturbation of the original  $x_\alpha^k$  and  $\leq$  so as to obtain a parametrized family of similar objects, on which probabilistic statements can later be made. The parameter space will be

$$\Omega = \left( \{0, 1, \dots, L\} \times \{1, \dots, M\} \right)^\mathbb{Z},$$

with a typical point denoted by  $\omega = (\omega_k)_{k \in \mathbb{Z}}$ , where  $\omega_k = (\ell_k, m_k) \in \{0, 1, \dots, L\} \times \{1, \dots, M\}$ .

The *new dyadic points*  $z_\alpha^k = z_\alpha^k(\omega_k)$  are defined by

$$z_\alpha^k := \begin{cases} x_\beta^{k+1} & \text{if } \text{label}_1(k, \alpha) = \ell_k, \text{ and } (k+1, \beta) \leq (k, \alpha), \text{ and } \text{label}_2(k+1, \beta) = m_k, \\ x_\alpha^k & \text{if } \text{label}_1(k, \alpha) \neq \ell_k, \text{ or } \nexists (k+1, \beta) \leq (k, \alpha) \text{ such that } \text{label}_2(k+1, \beta) = m_k. \end{cases}$$

A key feature of this definition is the following probabilistic statement, when  $\Omega$  is equipped with the natural probability measure  $\mathbb{P}_\omega$ , which makes all coordinates  $\omega_k = (\ell_k, m_k)$  are independent of each other and uniformly distributed over the finite set  $\{0, 1, \dots, L\} \times \{1, \dots, M\}$  (in other words,  $\ell_k$  is uniformly distributed over  $\{0, 1, \dots, L\}$ , and  $m_k$  over  $\{1, \dots, M\}$ , independently).

**2.3. Lemma.** *Let  $(k+1, \beta) \leq (k, \alpha)$  be fixed. Then*

$$\mathbb{P}_\omega(z_\alpha^k = x_\beta^{k+1}) \geq \frac{1}{(L+1)M}.$$

In other words, every old point on the level  $k+1$  has a positive (and bounded from below) probability of being a new point on the level  $k$ .

*Proof.* A sufficient condition that  $z_\alpha^k = x_\beta^{k+1}$  is that  $\ell_k = \text{label}_1(k, \alpha)$  (which has probability  $1/(L+1)$ ) and that  $m_k = \text{label}_2(k+1, \beta)$  (which has probability  $1/M$ ). Multiplying the probabilities of these independent events gives the claim.  $\square$

The point of the next result is that the new points behave qualitatively like the reference points, only with slightly weaker constants.

**2.4. Lemma.** *The new points satisfy*

$$d(z_\alpha^k, z_\beta^k) \geq \frac{1}{2A_0} \delta^k, \quad \min_\alpha d(x, z_\alpha^k) < 4A_0^2 \delta^k.$$

*Proof.* Consider the second bound first. Note that, in either of the two possibilities for the new point, we have  $z_\alpha^k = x_\beta^{k+1}$  for some  $(k+1, \beta) \leq (k, \alpha)$ ; in particular,  $d(x_\alpha^k, z_\alpha^k) < 2A_0 \delta^k$ , and thus

$$\min_\alpha d(x, z_\alpha^k) < \min_\alpha [A_0 d(x, x_\alpha^k) + A_0 d(x_\alpha^k, z_\alpha^k)] < 4A_0^2 \delta^k.$$

Let us then estimate  $d(z_\alpha^k, z_\beta^k)$  from below for  $\alpha \neq \beta$ . If  $(k, \alpha)$  and  $(k, \beta)$  are not neighbours, then by definition this distance is at least  $(2A_0)^{-1}\delta^k$ . So suppose that these pairs are neighbours.



Then they have different primary labels, and hence at least one of the new dyadic points, say  $z_\alpha^k$ , must satisfy  $z_\alpha^k = x_\alpha^k$ . On the other hand, if  $(k+1, \eta) \leq (k, \beta)$ , then  $(k+1, \eta) \not\leq (k, \alpha)$ , and we thus know that  $d(x_\eta^{k+1}, x_\alpha^k) \geq (2A_0)^{-1}\delta^k$ . But  $z_\beta^k$  will be one of these points  $x_\eta^{k+1}$ ; thus  $d(z_\beta^k, z_\alpha^k) \geq (2A_0)^{-1}\delta^k$ , as we claimed.  $\square$

The *new partial order*  $\leq_\omega$ ,  $\omega = (\omega_k)_{k \in \mathbb{Z}}$ , is set up as follows. We declare that

$$(2.5) \quad (k+1, \beta) \leq_\omega (k, \alpha) \quad \stackrel{\text{def}}{\iff} \quad \begin{cases} d(x_\beta^{k+1}, z_\alpha^k) < \frac{1}{4}A_0^{-2}\delta^k, & \text{or} \\ (k+1, \beta) \leq (k, \alpha) \text{ and } \nexists \gamma : d(x_\beta^{k+1}, z_\gamma^k) < \frac{1}{4}A_0^{-2}\delta^k. \end{cases}$$

In other words, to find the new parent of  $(k+1, \beta)$  for the new partial order  $\leq_\omega$ , we first check whether the reference point  $x_\beta^{k+1}$  is close (within distance  $\frac{1}{4}A_0^{-2}\delta^k$ ) to some new dyadic point  $z_\alpha^k$ . If yes, then the corresponding  $(k, \alpha)$  will be the new parent of  $(k+1, \beta)$ . If no such close point exists, then we simply use the original partial order  $\leq$  to decide the parent of  $(k+1, \beta)$ .

### Properties of the new points and order.

**2.6. Lemma.** *For any given  $k, \alpha, \beta$ , the truth or falsity of the relation  $(k+1, \beta) \leq_\omega (k, \alpha)$  depends only on the component  $\omega_k$  of  $\omega$ .*

*Proof.* In the defining conditions on the right of (2.5), the only dependence on  $\omega$  is via the new dyadic points  $z_\alpha^k, z_\gamma^k$ , and they depend only on  $\omega_k$  by definition.  $\square$

This explicit definition of  $\leq_\omega$  in terms of the geometric configuration of the points and the original partial order  $\leq$  is a novelty of our construction, where [19, 21] required a condition similar to (2.2), which only specifies the relation up to certain degrees of freedom. For us, a condition analogous to (2.2) is a *consequence* of the definition:

### 2.7. Lemma.

$$d(z_\beta^{k+1}, z_\alpha^k) < \frac{1}{5}A_0^{-3}\delta^k \implies (k+1, \beta) \leq_\omega (k, \alpha) \implies d(z_\beta^{k+1}, z_\alpha^k) < 5A_0^3\delta^k.$$

*Proof.* Let first  $d(z_\beta^{k+1}, z_\alpha^k) < \frac{1}{5}A_0^{-3}\delta^k$ . Then

$$d(x_\beta^{k+1}, z_\alpha^k) \leq A_0 d(x_\beta^{k+1}, z_\beta^{k+1}) + A_0 d(z_\beta^{k+1}, z_\alpha^k) < 2A_0^2\delta^{k+1} + \frac{1}{5}A_0^{-2}\delta^k \leq \frac{1}{4}A_0^{-2}\delta^k,$$

and hence  $(k+1, \beta) \leq_\omega (k, \alpha)$  by definition.

Let then  $(k+1, \beta) \leq_\omega (k, \alpha)$ . If  $d(x_\beta^{k+1}, z_\alpha^k) < \frac{1}{4}A_0^{-2}\delta^k$ , then

$$d(z_\beta^{k+1}, z_\alpha^k) \leq A_0 d(z_\beta^{k+1}, x_\beta^{k+1}) + A_0 d(x_\beta^{k+1}, z_\alpha^k) < 2A_0^2\delta^{k+1} + \frac{1}{4}A_0^{-1}\delta^k < 5A_0^3\delta^k.$$

Otherwise, we have  $(k+1, \beta) \leq (k, \alpha)$ , and hence

$$\begin{aligned} d(z_\beta^{k+1}, z_\alpha^k) &\leq A_0^2 d(z_\beta^{k+1}, x_\beta^{k+1}) + A_0^2 d(x_\beta^{k+1}, x_\alpha^k) + A_0 d(x_\alpha^k, z_\alpha^k) \\ &< 2A_0^3\delta^{k+1} + 2A_0^3\delta^k + 2A_0^2\delta^k < 5A_0^3\delta^k. \end{aligned}$$

$\square$

We can iterate this as follows:

### 2.8. Lemma.

$$d(z_\beta^\ell, z_\alpha^k) < \frac{1}{6}A_0^{-4}\delta^k \implies (\ell, \beta) \leq_\omega (k, \alpha) \implies d(z_\beta^\ell, z_\alpha^k) < 6A_0^4\delta^k.$$

*Proof.* The second implication follows from the second implication of Lemma 2.7 with the triangle inequality:

$$(\ell, \beta) \leq_\omega (k, \alpha) \implies d(z_\beta^\ell, z_\alpha^k) < \sum_{j=0}^{\ell-k-1} 5A_0^3\delta^{k+j} \cdot A_0^{j+1} < \frac{5A_0^4\delta^k}{1-A_0\delta} < 6A_0^4\delta^k.$$

For the first implication, if  $\ell = k$ , then the closeness of the points implies that  $\beta = \alpha$ . If  $\ell > k$ , consider  $\gamma$  such that  $(\ell, \beta) \leq_\omega (k+1, \gamma)$ . By what we just proved,  $d(z_\beta^\ell, z_\gamma^{k+1}) < 6A_0^4\delta^{k+1}$ , and hence

$$d(z_\gamma^{k+1}, z_\alpha^k) \leq A_0 \left( \frac{1}{6}A_0^{-4}\delta^k + 6A_0^4\delta^{k+1} \right) < \frac{1}{5}A_0^{-3}\delta^k;$$

thus  $(\ell, \beta) \leq_\omega (k+1, \gamma) \leq_\omega (k, \alpha)$ , where the last step follows from Lemma 2.7.  $\square$

**Dyadic cubes.** With the auxiliary objects at hand, the **dyadic cubes** are easy to define. As in [19], we introduce three families of these cubes—the preliminary, the closed, and the open:

$$\hat{Q}_\alpha^k(\omega) := \{z_\beta^\ell(\omega) : (\ell, \beta) \leq_\omega (k, \alpha)\},$$

$$\bar{Q}_\alpha^k(\omega) := \overline{\hat{Q}_\alpha^k(\omega)}, \quad \tilde{Q}_\alpha^k(\omega) := \text{interior } \bar{Q}_\alpha^k(\omega).$$

Note that  $\hat{Q}_\alpha^k(\omega)$ , and hence  $\bar{Q}_\alpha^k(\omega)$  and  $\tilde{Q}_\alpha^k(\omega)$ , only depends on  $\omega_\ell$  for  $\ell \geq k$ .

The rest of the section is concerned with the properties of these cubes. We first deal with several properties valid for an arbitrary fixed choice of the parameter  $\omega \in \Omega$ , and finally present a probabilistic statement concerning a random choice of  $\omega$  under the natural product probability measure on the set  $\Omega$ .

**2.9. Lemma.**

$$\bar{Q}_\alpha^k(\omega) \subseteq B(z_\alpha^k, 6A_0^4\delta^k).$$

*Proof.* Let  $x \in \bar{Q}_\alpha^k(\omega)$ . Then  $x$  is a limit of some points  $z_\beta^\ell$  with  $(\ell, \beta) \leq_\omega (k, \alpha)$ , and we may choose a subsequence with  $(\ell, \beta) \leq_\omega (k+1, \gamma) \leq_\omega (k, \alpha)$  for some  $\gamma$ . Then

$$d(z_\alpha^k, x) \leq A_0 d(z_\alpha^k, z_\gamma^{k+1}) + A_0^2 d(z_\gamma^{k+1}, z_\beta^\ell) + A_0^2 d(z_\beta^\ell, x),$$

where the sum of the first two terms, by Lemma 2.7 and (2) is  $< 5A_0^4\delta^k + 6A_0^6\delta^{k+1} < 6A_0^4\delta^k$ , and the last term becomes arbitrarily small.  $\square$

**2.10. Lemma.** *We have the following covering properties for each fixed  $k \in \mathbb{Z}$ :*

$$X = \bigcup_{\alpha} \bar{Q}_\alpha^k(\omega), \quad \bar{Q}_\alpha^k(\omega) = \bigcup_{\beta: (k+1, \beta) \leq_\omega (k, \alpha)} \bar{Q}_\beta^{k+1}(\omega).$$

*Proof.* Every  $x \in X$  is the limit of points  $z_{\theta(m)}^m$  where  $m \rightarrow \infty$ . When  $m \geq k$ , we have  $(m, \theta(m)) \leq (k, \alpha)$  for some  $\alpha$ , and hence  $\bigcup_{\alpha} \bar{Q}_\alpha^k(\omega)$  is dense in  $X$ . On the other hand, the uniform separation of the centres  $z_\alpha^k$ , the uniformly bounded diameter of the sets  $\bar{Q}_\alpha^k(\omega)$ , and the geometric doubling property imply that this union is locally finite. Hence the closure of the union is the union of the closures.

Similarly, it is immediate that  $\bar{Q}_\alpha^k(\omega) = \bigcup_{\beta: (k+1, \beta) \leq_\omega (k, \alpha)} \bar{Q}_\beta^{k+1}(\omega)$ , and the second claim follows by taking the closures, observing that the union is finite.  $\square$

**2.11. Lemma.** *The closed and open cubes of the same generation are disjoint:*

$$\bar{Q}_\alpha^k(\omega) \cap \tilde{Q}_\beta^k(\omega) = \emptyset \quad (\alpha \neq \beta).$$

*In fact,*

$$\tilde{Q}_\alpha^k(\omega)^c = \bigcup_{\beta \neq \alpha} \bar{Q}_\beta^k(\omega).$$

*Proof.* We first check the weaker statement that

$$\bar{Q}_\alpha^k(\omega) \cap \hat{Q}_\beta^k(\omega) = \emptyset \quad (\alpha \neq \beta).$$

Indeed, for contradiction, let  $x \in \bar{Q}_\alpha^k(\omega) \cap \hat{Q}_\beta^k(\omega)$ . Thus  $x = z_\gamma^\ell$  with  $(\ell, \gamma) \leq_\omega (k, \beta)$ , and also  $x = \lim_{m \rightarrow \infty} z_{\theta(m)}^m$  with  $(m, \theta(m)) \leq (k, \alpha)$ . For  $m$  large enough, we deduce that  $d(z_{\theta(m)}^m, z_\gamma^\ell) < \frac{1}{6}A_0^{-4}\delta^\ell$  and thus, by Lemma 2.8,  $(m, \theta(m)) \leq_\omega (\ell, \gamma) \leq_\omega (k, \beta)$ , a contradiction.

To prove the actual first claim, again by contradiction, let  $x \in \bar{Q}_\alpha^k(\omega) \cap \tilde{Q}_\beta^k(\omega)$ . Hence  $x = \lim_{m \rightarrow \infty} x_m$  with  $x_m \in \hat{Q}_\alpha^k(\omega)$ . Since  $\tilde{Q}_\beta^k(\omega)$  is open by definition and  $x_m \rightarrow x$ , it follows that  $x_m \in \bar{Q}_\beta^k(\omega) \subseteq \tilde{Q}_\beta^k(\omega)$  for large enough  $m$ . So in fact  $x_m \in \hat{Q}_\alpha^k(\omega) \cap \bar{Q}_\beta^k(\omega)$ , but this is empty by the first part of the proof. The claim follows.

For the second claim, it is immediate from the first claim that  $\tilde{Q}_\alpha^k(\omega) \subseteq \bigcap_{\beta \neq \alpha} \bar{Q}_\beta^k(\omega)^c = \left( \bigcup_{\beta \neq \alpha} \bar{Q}_\beta^k(\omega) \right)^c =: \mathcal{O}_\alpha^k$ . On the other hand, by local finiteness, the union defining  $(\mathcal{O}_\alpha^k)^c$  is closed, and therefore  $\mathcal{O}_\alpha^k$  itself, open. Since  $X = \bar{Q}_\alpha^k(\omega) \cup (\mathcal{O}_\alpha^k)^c$ , we have  $\mathcal{O}_\alpha^k \subseteq \bar{Q}_\alpha^k(\omega)$ , and since  $\tilde{Q}_\alpha^k(\omega)$  is the largest open set with this property, we get  $\mathcal{O}_\alpha^k \subseteq \tilde{Q}_\alpha^k(\omega)$ .  $\square$



**2.12. Lemma.**

$$B(z_\alpha^k, \frac{1}{6}A_0^{-5}\delta^k) \subseteq \tilde{Q}_\alpha^k(\omega).$$

*Proof.* It suffices to prove that  $B(z_\alpha^k, \frac{1}{6}A_0^{-5}\delta^k)$  is disjoint from each  $\tilde{Q}_\beta^k(\omega)$  with  $\beta \neq \alpha$ . For contradiction, let  $x \in B(z_\alpha^k, \frac{1}{6}A_0^{-5}\delta^k) \cap \tilde{Q}_\beta^k(\omega)$ . Thus  $x = \lim_{m \rightarrow \infty} z_{\theta(m)}^m$  for some  $(m, \theta(m)) \leq_\omega (k, \beta)$ . Then

$$d(z_{\theta(m)}^m, z_\alpha^k) \leq A_0 d(z_{\theta(m)}^m, x) + A_0 d(x, z_\alpha^k) < \frac{1}{6}A_0^{-4}\delta^k$$

for large enough  $m$ , since the second term is strictly smaller than this bound, and the first term tends to zero as  $m \rightarrow \infty$ . But then Lemma 2.8 says that  $(m, \theta(m)) \leq_\omega (k, \alpha)$ , a contradiction with  $(m, \theta(m)) \leq_\omega (k, \beta)$ .  $\square$

The following theorem summarizes the above properties of the dyadic cubes for a fixed parameter  $\omega$ , and supplements the key statement about their probabilistic behaviour under the random choice of  $\omega \in \Omega$ .

**2.13. Theorem.** *For any fixed  $\omega \in \Omega := (\{0, 1, \dots, L\} \times \{1, \dots, M\})^{\mathbb{Z}}$ , the cubes satisfy the following relations: the covering properties*

$$X = \bigcup_{\alpha} \tilde{Q}_\alpha^k(\omega), \quad \tilde{Q}_\alpha^k(\omega) = \bigcup_{\beta: (k+1, \beta) \leq_\omega (k, \alpha)} \tilde{Q}_\beta^{k+1}(\omega);$$

*the mutual disjointness property*

$$\tilde{Q}_\alpha^k(\omega) \cap \tilde{Q}_\beta^k(\omega) = \emptyset \quad (\alpha \neq \beta);$$

*and the comparability with balls:*

$$B(z_\alpha^k, \frac{1}{6}A_0^{-5}\delta^k) \subseteq \tilde{Q}_\alpha^k(\omega) \subseteq \bar{Q}_\alpha^k(\omega) \subseteq B(z_\alpha^k, 6A_0^4\delta^k).$$

Moreover, when  $\Omega$  is equipped with the natural probability measure  $\mathbb{P}_\omega$ , we have for some  $\eta \in (0, 1]$  the small boundary layer property:

$$(2.14) \quad \mathbb{P}_\omega \left( x \in \bigcup_{\alpha} \partial_\varepsilon Q_\alpha^k(\omega) \right) \leq C\varepsilon^\eta \quad \left( \partial_\varepsilon Q_\alpha^k(\omega) := \{y \in \bar{Q}_\alpha^k(\omega) : d(y, {}^c\tilde{Q}_\alpha^k(\omega)) < \varepsilon\delta^k\} \right);$$

*and in particular the negligible boundary property:*

$$\mathbb{P}_\omega \left( x \in \bigcup_{k, \alpha} \partial Q_\alpha^k(\omega) \right) = 0 \quad \left( \partial Q_\alpha^k(\omega) := \bar{Q}_\alpha^k(\omega) \setminus \tilde{Q}_\alpha^k(\omega) \right).$$

**Henceforth,  $\eta$  will always designate the fixed positive constant provided by this theorem.** It will reappear as the Hölder-regularity index of our splines and wavelets.

*Proof.* It only remains to check the probabilistic statements.

Let  $x \in X$ ,  $k \in \mathbb{Z}$ , and  $\varepsilon > 0$  be fixed. For every  $\ell = k, k+1, \dots$ , there is some  $\gamma$  so that  $d(x, x_\gamma^{\ell+1}) < 2A_0\delta^{\ell+1}$ . Now suppose that for some  $\ell$ , the point  $x_\gamma^{\ell+1}$  is chosen as a new dyadic point  $z_\beta^\ell$ , and recall that  $B(z_\beta^\ell, \frac{1}{6}A_0^{-5}\delta^\ell) \subseteq \tilde{Q}_\beta^\ell(\omega)$ . Thus,

$$d(x, {}^c\tilde{Q}_\beta^\ell(\omega)) \geq \frac{1}{A_0}d(z_\beta^\ell, {}^c\tilde{Q}_\beta^\ell(\omega)) - d(x, z_\beta^\ell) \geq \left(\frac{1}{6}A_0^{-6} - 2A_0\delta\right)\delta^\ell \geq \frac{1}{7}A_0^{-6}\delta^\ell \geq \varepsilon\delta^k$$

provided that  $\ell \leq k + \log(7A_0^6\varepsilon)/\log\delta$ . In particular, we have  $x \in \tilde{Q}_\beta^\ell(\omega) \subseteq \tilde{Q}_\alpha^k(\omega)$  if  $(\ell, \beta) \leq_\omega (k, \alpha)$ , while  $d(x, {}^c\tilde{Q}_\alpha^k(\omega)) \geq d(x, {}^c\tilde{Q}_\beta^\ell(\omega)) \geq \varepsilon\delta^k$ . Thus  $x$  cannot be in  $\bigcup_{\alpha} \partial_\varepsilon Q_\alpha^k(\omega)$  in this case.

In other words, in order that  $x$  belongs to the union  $\bigcup_{\alpha} \partial_\varepsilon Q_\alpha^k(\omega)$ , it is necessary that *none* of the points  $x_\gamma^{\ell+1}$ , where  $k \leq \ell \leq k + \log(7A_0^6\varepsilon)/\log\delta$ , is chosen as a new dyadic point. But, by Lemma 2.3, every  $x_\gamma^{\ell+1}$  has a probability at least  $\tau := \frac{1}{M(L+1)}$  of being chosen. So the probability that  $x_\gamma^{\ell+1}$  is not chosen is at most  $1 - \tau$ . Moreover, these events for different levels  $\ell$  are independent of each other. Hence the probability that none of the  $x_\gamma^{\ell+1}$  is chosen, for  $k \leq \ell \leq k + \log(7A_0^6\varepsilon)/\log\delta$ , is at most

$$(1 - \tau)^{\log(7A_0^6\varepsilon)/\log\delta} = (7A_0^6\varepsilon)^{\log(1-\tau)/\log\delta} = C\varepsilon^\eta,$$

where  $\eta := \log(1 - \tau)/\log \delta > 0$  (since both  $\delta, 1 - \tau \in (0, 1)$ ) and  $C = (7A_0^6)^\eta$ . This is exactly as claimed.

The negligible boundary property follows from the small boundary property as  $\varepsilon \rightarrow 0$ .  $\square$

We conclude this section with the observation that in our construction, the original dyadic point  $x_\alpha^k$  may also be viewed as a ‘centre’ of the new dyadic cubes  $\bar{Q}_\alpha^k(\omega)$  for all  $\omega \in \Omega$ :

**2.15. Lemma.**  $B(x_\alpha^k, \frac{1}{8}A_0^{-3}\delta^k) \subseteq \bar{Q}_\alpha^k(\omega) \subseteq \bar{B}(x_\alpha^k, 8A_0^5\delta^k)$  for any  $\omega \in \Omega$ .

*Proof.* By Lemma 2.9, we have for any  $x \in \bar{Q}_\alpha^k$  that

$$d(x, x_\alpha^k) \leq A_0 d(x, z_\alpha^k) + A_0 d(z_\alpha^k, x_\alpha^k) < 6A_0^5\delta^k + 2A_0^2\delta^k \leq 8A_0^5\delta^k,$$

which shows that  $\bar{Q}_\alpha^k \subseteq B(x_\alpha^k, 8A_0^5\delta^k)$ .

For the other inclusion, let  $d(x, x_\alpha^k) < \frac{1}{8}A_0^{-3}\delta^k$ . Since  $X = \bigcup_\eta \bar{Q}_\eta^{k+1}(\omega)$ , we have  $x \in \bar{Q}_\eta^{k+1}(\omega) \subseteq B(x_\eta^{k+1}, 8A_0^5\delta^{k+1})$  for some  $\eta$ . But then

$$d(x_\eta^{k+1}, x_\alpha^k) \leq A_0 d(x_\eta^{k+1}, x) + A_0 d(x, x_\alpha^k) < 8A_0^6\delta^{k+1} + \frac{1}{8}A_0^{-2}\delta^k < \frac{1}{2}A_0^{-1}\delta^k.$$

By another application of (2.2), this implies that  $(k+1, \eta) \leq (k, \alpha)$ .

We want to prove that in fact  $(k+1, \eta) \leq_\omega (k, \alpha)$ , since then  $x \in \bar{Q}_\eta^{k+1}(\omega) \subseteq \bar{Q}_\alpha^k(\omega)$ . By (2.5), the only potential obstacle to this is that  $d(x_\eta^{k+1}, z_\gamma^k) < \frac{1}{4}A_0^{-2}\delta^k$  for some  $\gamma \neq \alpha$ . So suppose this is the case, and recall that  $z_\gamma^k = x_\theta^{k+1}$  for some  $(k+1, \theta) \leq (k, \gamma)$ . Then

$$\begin{aligned} d(x_\theta^{k+1}, x_\alpha^k) &\leq A_0 d(z_\gamma^k, x_\eta^{k+1}) + A_0^2 d(x_\eta^{k+1}, x) + A_0^2 d(x, x_\alpha^k) \\ &< \frac{1}{4}A_0^{-1}\delta^k + 8A_0^7\delta^{k+1} + \frac{1}{8}A_0^{-1}\delta^k < \frac{1}{2}A_0^{-1}\delta^k, \end{aligned}$$

and hence  $(k+1, \theta) \leq (k, \alpha)$  by (2.2). But also  $(k+1, \theta) \leq (k, \gamma)$ , hence  $\gamma = \alpha$ , a contradiction.

We have shown that an arbitrary  $x \in B(x_\alpha^k, \frac{1}{8}A_0^{-3}\delta^k)$  satisfies  $x \in \bar{Q}_\eta^{k+1}(\omega)$  for some  $(k+1, \eta) \leq_\omega (k, \alpha)$ ; thus  $x \in \bar{Q}_\alpha^k(\omega)$ .  $\square$

### 3. CONSTRUCTION OF SPLINES

The construction of splines on  $X$ , and the proof of their basic properties, is now an easy consequence of the preparations from the previous section. For every  $(k, \alpha)$ , we define the **spline function**

$$(3.1) \quad s_\alpha^k(x) := \mathbb{P}_\omega(x \in \bar{Q}_\alpha^k(\omega)).$$

**3.2. Theorem.** *The splines (3.1) satisfy the following properties: bounded support*

$$(3.3) \quad 1_{B(x_\alpha^k, \frac{1}{8}A_0^{-3}\delta^k)}(x) \leq s_\alpha^k(x) \leq 1_{B(x_\alpha^k, 8A_0^5\delta^k)}(x);$$

*the interpolation and reproducing properties*

$$(3.4) \quad s_\alpha^k(x_\beta^k) = \delta_{\alpha\beta}, \quad \sum_\alpha s_\alpha^k(x) = 1, \quad s_\alpha^k(x) = \sum_\beta p_{\alpha\beta}^k \cdot s_\beta^{k+1}(x)$$

where  $\{p_{\alpha\beta}^k\}_\beta$  is a finitely nonzero set of nonnegative coefficients with  $\sum_\beta p_{\alpha\beta}^k = 1$ ; and Hölder-continuity

$$|s_\alpha^k(x) - s_\alpha^k(y)| \leq C \left( \frac{d(x, y)}{\delta^k} \right)^\eta.$$

Regarding (3.3), remark that it is a bit unusual that a spline be a non-zero constant on part of its support.

*Proof.* The relations (3.3) are immediate from Lemma 2.15, and this implies in particular that  $s_\alpha^k(x_\alpha^k) = 1$ .

Since the boundaries have vanishing probability and  $X = \bigcup_\alpha \bar{Q}_\alpha^k(\omega)$ , it follows that

$$(3.5) \quad \sum_\alpha s_\alpha^k(x) = \mathbb{P}_\omega(x \in X) = 1.$$

Since these functions are nonnegative and  $s_\alpha^k(x_\alpha^k) = 1$ , it must be that  $s_\alpha^k(x_\beta^k) = 0$  for  $\beta \neq \alpha$ , and hence in fact one has the **interpolation property**

$$(3.6) \quad s_\alpha^k(x_\beta^k) = \delta_{\alpha\beta}.$$

From the identity

$$\bar{Q}_\alpha^k(\omega) = \bigcup_{\beta: (k+1, \beta) \leq_\omega (k, \alpha)} \bar{Q}_\beta^{k+1}(\omega),$$

(and using again the vanishing probability of the boundaries) we also have

$$\begin{aligned} s_\alpha^k(x) &= \mathbb{P}_\omega \left( x \in \bigcup_{\beta: (k+1, \beta) \leq_\omega (k, \alpha)} \bar{Q}_\beta^{k+1}(\omega) \right) \\ &= \sum_\beta \mathbb{P}_\omega \left( \{ (k+1, \beta) \leq_\omega (k, \alpha) \} \cap \{ x \in \bar{Q}_\beta^{k+1}(\omega) \} \right) \\ &= \sum_\beta \mathbb{P}_\omega \left( (k+1, \beta) \leq_\omega (k, \alpha) \right) \mathbb{P}_\omega \left( x \in \bar{Q}_\beta^{k+1}(\omega) \right) \\ &= \sum_\beta \mathbb{P}_\omega \left( (k+1, \beta) \leq_\omega (k, \alpha) \right) s_\beta^{k+1}(x) =: \sum_\beta p_{\alpha\beta}^k \cdot s_\beta^{k+1}(x), \end{aligned}$$

where the key third step used the independence of the two events; namely, the event  $(k+1, \beta) \leq_\omega (k, \alpha)$  depends only on  $\omega_k$ , while the cube  $\bar{Q}_\beta^{k+1}(\omega)$  depends on  $\omega_\ell$  for  $\ell \geq k+1$ . The support properties of the splines readily imply that only boundedly many of the coefficients  $p_{\alpha\beta}^k$  are nonzero for a given  $(k, \alpha)$ , so that  $\text{span}\{s_\alpha^k\}_\alpha \subseteq \text{span}\{s_\beta^{k+1}\}_\beta$ .

The **Hölder-continuity of the splines** follows from the probabilistic smallness of the boundary regions, as expressed by (2.14). Indeed,

$$\begin{aligned} |s_\alpha^k(x) - s_\alpha^k(y)| &= \left| \mathbb{P}_\omega \left( x \in \bar{Q}_\alpha^k(\omega) \right) - \mathbb{P}_\omega \left( y \in \bar{Q}_\alpha^k(\omega) \right) \right| \\ &= \left| \int_\Omega (1_{\omega: x \in \bar{Q}_\alpha^k(\omega)} - 1_{\omega: y \in \bar{Q}_\alpha^k(\omega)}) d\mathbb{P}_\omega \right| \\ &\leq \int_\Omega (1_{\omega: x \in \bar{Q}_\alpha^k(\omega), y \notin \bar{Q}_\alpha^k(\omega)} + 1_{\omega: y \in \bar{Q}_\alpha^k(\omega), x \notin \bar{Q}_\alpha^k(\omega)}) d\mathbb{P}_\omega \\ &= \mathbb{P}_\omega \left( x \in \bar{Q}_\alpha^k(\omega), y \notin \bar{Q}_\alpha^k(\omega) \right) + \mathbb{P}_\omega \left( y \in \bar{Q}_\alpha^k(\omega), x \notin \bar{Q}_\alpha^k(\omega) \right) \\ &\leq \mathbb{P}_\omega \left( x \in \partial_{d(x,y)\delta^{-k}} \bar{Q}_\alpha^k(\omega) \right) + \mathbb{P}_\omega \left( y \in \partial_{d(x,y)\delta^{-k}} \bar{Q}_\alpha^k(\omega) \right) \\ &\leq C \left( \frac{d(x,y)}{\delta^k} \right)^\eta. \end{aligned} \quad \square$$

#### 4. AUXILIARY RESULTS ON QUASI-METRIC SPACES

We include this intermediate section to collect some auxiliary material so as to streamline the subsequent analysis. This section is mostly concerned with difficulties of a quasi-metric in contrast to a metric. The results are mostly part of the folklore, but somewhat difficult to find in the literature in full generality, since additional assumptions on the space are usually imposed. (After completion of this work, we became aware of the recent preprint [31] where the density result below is also proved in full generality by a completely different argument.)

For the beginning of this section, as above, we only assume that  $X$  is a geometrically doubling quasi-metric space. **Throughout, let  $\eta > 0$  denote the fixed positive constant from the small boundary property of the dyadic cubes and the Hölder regularity of the splines.**

**4.1. Lemma.** *Let  $F \subseteq G \subseteq X$  be sets with*

$$d(F, G^c) := \inf_{x \in F, y \notin G} d(x, y) = \Delta > 0.$$

Then there exists a function  $\varphi : X \rightarrow \mathbb{R}$  with  $1_F \leq \varphi \leq 1_G$  and

$$|\varphi(x) - \varphi(y)| \leq C \left( \frac{d(x, y)}{\Delta} \right)^\eta.$$

*Proof.* Let  $k$  be the smallest integer so that  $16A_0^6\delta^k \leq \Delta$ . We set

$$\varphi := \sum_{\alpha: B(x_\alpha^k, 8A_0^5\delta^k) \cap F \neq \emptyset} s_\alpha^k.$$

Then  $\varphi(x) = 1$  for  $x \in F$ , since  $\sum_\alpha s_\alpha^k(x) = 1$ , and the sum defining  $\varphi(x)$  contains all  $s_\alpha^k(x)$  whose support intersects  $F$ . To prove that  $\varphi(z) = 0$  for  $z \in G^c$ , we need to check that no  $\text{supp } s_\alpha^k \subseteq B(x_\alpha^k, 8A_0^5\delta^k)$  appearing in  $\varphi$  can intersect  $G^c$ . To this end, let  $y \in B(x_\alpha^k, 8A_0^5\delta^k) \cap F$  (which exists by definition) and  $z \in B(x_\alpha^k, 8A_0^5\delta^k) \cap G^c$  (whose assumed existence should lead to a contradiction). Then

$$d(F, G^c) \leq d(y, z) < A_0 d(y, x_\alpha^k) + A_0 d(x_\alpha^k, z) \leq 2A_0 \cdot 8A_0^5\delta^k \leq \Delta,$$

a contradiction indeed. So we have  $1_F \leq \varphi \leq 1_G$ , and it remains to check the Hölder-continuity. By geometric doubling, there are only boundedly many  $s_\alpha^k$ , whose support contains either  $x$  or  $y$ . Each of these  $s_\alpha^k$  satisfy the estimate

$$|s_\alpha^k(x) - s_\alpha^k(y)| \leq C \left( \frac{d(x, y)}{\delta^k} \right)^\eta \leq C \left( \frac{d(x, y)}{\Delta} \right)^\eta,$$

and hence so does their sum over boundedly many indices  $\alpha$ . This completes the proof.  $\square$

**4.2. Corollary.** *For every  $B(x, r)$ , there exists a function  $\varphi : X \rightarrow \mathbb{R}$  with  $1_{B(x, r)} \leq \varphi \leq 1_{B(x, 2A_0 r)}$  and*

$$|\varphi(x) - \varphi(y)| \leq C \left( \frac{d(x, y)}{r} \right)^\eta.$$

*Proof.* If  $y \in B(x, r)$  and  $z \in B(x, 2A_0 r)^c$ , then

$$d(y, z) \geq \frac{1}{A_0} d(z, x) - d(x, y) > 2r - r = r;$$

thus  $d(B(x, r), B(x, 2A_0 r)^c) \geq r$ , and the previous lemma applies.  $\square$

We formulate a quasi-metric version of a well-known covering lemma for metric spaces, cf. [18, Theorem 1.2]; the quasi-metric extension is obtained *mutatis mutandis*.

**4.3. Lemma.** *Let  $\mathcal{B}$  be a family of balls  $B(x, r)$  in  $X$ , with bounded radii and contained in a bounded set. Then there exists a pairwise disjoint subcollection  $\{B_i\}_{i=1}^\infty$ , whose concentric expansions  $5A_0^2 B_i$  cover all original  $B(x, r) \in \mathcal{B}$ .*

**4.4. Lemma.**  *$X$  is a countable union of bounded open sets.*

*Proof.* Recall that the interior of a set  $E$  is the open set  $(\overline{E^c})^c \subseteq E$ . It is easy to check that

$$B(x, r/A_0) \subseteq \text{interior } B(x, r);$$

hence, with any base point  $x_0 \in X$ , the bounded open sets  $\text{interior } B(x_0, n)$ ,  $n \in \mathbb{N}$ , provide the required covering.  $\square$

So far everything has been based on the space geometry, i.e., the properties of the quasi-distance only and geometric doubling. We next add a measure  $\mu$  into our considerations.

**4.5. Proposition.** *Let  $\mu$  be a non-trivial Borel measure on  $X$ , finite on bounded Borel sets. Let  $1 \leq p < \infty$ . Then Hölder- $\eta$ -continuous functions of bounded support are dense in  $L^p(\mu)$ , where  $\eta$  is the Hölder exponent of the splines.*

*Proof.* By the density of simple functions due to general measure theory, it suffices to show that for every bounded Borel set  $E$  and every  $\epsilon > 0$ , there exists a boundedly supported Hölder- $\eta$ -continuous function  $\varphi$  with  $\|1_E - \varphi\|_p < 2\epsilon$ . By a general result concerning Borel measures [12, Theorem 2.2.2 (ii)]<sup>1</sup>, there is an open set  $G \supseteq E$  such that  $\mu(G \setminus E) < \epsilon^p$ . For every  $x \in G$ , we choose a radius  $r_x$  small enough so that  $B(x, 4A_0^2 r_x) \subseteq G$ . The balls  $B(x, 5^{-1}A_0^{-2}r_x)$  cover  $G$ , and hence there are pairwise disjoint balls  $B(x_i, 5^{-1}A_0^{-2}r_i)$ ,  $r_i := r_{x_i}$ , so that  $B(x_i, r_i)$  cover  $G$ . As we have  $B(x, r) \subseteq \overline{B(x, r)} \subseteq B(x, 2A_0 r)$  for any ball, with  $\overline{B}$  being the topological closure of the ball  $B$ ,

$$G \subseteq \bigcup_{i=1}^{\infty} \overline{B(x_i, r_i)} \subseteq \bigcup_{i=1}^{\infty} B(x_i, 2A_0 r_i) \subseteq G.$$

Thus

$$\mu(G) = \lim_{n \rightarrow \infty} \mu\left(\bigcup_{i=1}^n \overline{B(x_i, r_i)}\right) =: \lim_{n \rightarrow \infty} \mu(G_n).$$

Let us fix  $n$  so large that  $\mu(G \setminus G_n) \leq \epsilon^p$ . We check that  $d(G_n, G^c) > 0$ . Indeed, if  $x \in G_n$ , then  $x \in \overline{B(x_i, r_i)} \subseteq B(x_i, 2A_0 r_i)$  for some  $i = 1, \dots, n$ , whereas  $B(x_i, 4A_0^2 r_i) \subseteq G$ . Hence

$$d(G_n, G^c) \geq \min_{i=1, \dots, n} d(B(x_i, 2A_0 r_i), B(x_i, 4A_0^2 r_i)^c) \geq \min_{i=1, \dots, n} r_i > 0.$$

By Lemma 4.1, there exists a Hölder- $\eta$ -continuous  $\varphi$  with  $1_{G_n} \leq \varphi \leq 1_G$ . Hence  $\|1_G - \varphi\|_p \leq \|1_G - 1_{G_n}\|_p = \mu(G \setminus G_n)^{1/p} < \epsilon$ , and finally  $\|1_E - \varphi\|_p < 2\epsilon$ .  $\square$

## 5. $L^2$ THEORY AND MULTIREOLUTION ANALYSIS

We now return to the development of the spline theory in the presence of a nontrivial Borel measure  $\mu$  on  $(X, d)$ . From now on, we assume that  $(X, d, \mu)$  is a space of homogeneous type in the most general sense, namely that in addition to having a quasi-metric  $d$  we only assume the doubling condition which reads: for all  $x \in X$  and  $r > 0$

$$0 < \mu(B(x, 2r)) \leq C_\mu \cdot \mu(B(x, r)) < \infty.$$

This inequality makes sense if balls are Borel sets. But this may not be the case. However, their topological closures are (because they are closed sets by definition) and satisfy  $B(x, r) \subseteq \overline{B(x, r)} \subseteq B(x, 2A_0 r)$ . In that case, up to changing constants throughout, we may change balls to their closures in the above definition of the doubling condition, thus avoiding to resort to the outer measure associated to  $\mu$ . To simplify matters though, we assume that balls are Borel sets and leave to the reader the modifications if not.

The doubling condition on  $(X, d, \mu)$  implies that  $(X, d)$  is geometrically doubling (see [6]), so that the spline functions, built independently of measure considerations, are at our disposal. We now show that they provide a multiresolution analysis of  $L^2(\mu)$ , essentially in the sense of Sweldens [39, Definition 3.1], who considered the case of general measures on  $\mathbb{R}^n$ . This consist of all properties of a classical multiresolution analysis of Meyer [30, Definition 2.1], to the extent that this definition is meaningful in a quasi-metric space context: the classical postulates dealing with translations and dilations, specific to the Euclidean space and the Lebesgue measure, are now meaningless.

**5.1. Theorem.** *Let  $V_k$  be the closed linear span of  $\{s_\alpha^k\}_\alpha$  in  $L^2(\mu)$ . Then  $V_k \subseteq V_{k+1}$ , and*

$$\overline{\bigcup_{k \in \mathbb{Z}} V_k} = L^2(\mu), \quad \bigcap_{k \in \mathbb{Z}} V_k = \begin{cases} \{0\}, & \text{if } X \text{ is unbounded,} \\ V_{k_0} = \{\text{constants}\}, & \text{if } X \text{ is bounded,} \end{cases}$$

<sup>1</sup>This result is stated in a metric space with an outer measure but the argument is valid for any topological space for a Borel measure. Moreover, it is assumed that  $E$  can be covered by countably many open sets with bounded measure which is granted from Lemma 4.4 and our assumptions on  $\mu$ .

where  $k_0$  is some integer. Moreover, the functions  $s_\alpha^k/\sqrt{\mu_\alpha^k}$  form a Riesz basis of  $V_k$ : for all sequences of numbers  $\lambda_\alpha$ , we have the two-sided estimate

$$\left\| \sum_\alpha \lambda_\alpha s_\alpha^k \right\|_{L^2(\mu)} \approx \left( \sum_\alpha |\lambda_\alpha|^2 \mu_\alpha^k \right)^{1/2},$$

with  $\mu_\alpha^k := \mu(B(x_\alpha^k, \delta^k))$ .

*Proof.* The nesting property  $V_k \subseteq V_{k+1}$  is immediate from the reproducing properties of the splines.

*The Riesz basis property.* We use properties (3.3) and (3.4) of the splines. At any point  $x$ , the values  $s_\alpha^k(x)$  are nonnegative and sum up to 1 (when  $\alpha$  is the summation variable and  $k$  is fixed). Hence

$$\left| \sum_\alpha \lambda_\alpha s_\alpha^k(x) \right|^2 \leq \sum_\alpha |\lambda_\alpha|^2 s_\alpha^k(x) \leq \sum_\alpha |\lambda_\alpha|^2 1_{B(x_\alpha^k, 8A_0^5 \delta^k)}(x).$$

Thus

$$\int \left| \sum_\alpha \lambda_\alpha s_\alpha^k(x) \right|^2 d\mu(x) \leq \sum_\alpha |\lambda_\alpha|^2 \mu(B(x_\alpha^k, 8A_0^5 \delta^k)) \leq C \sum_\alpha |\lambda_\alpha|^2 \mu(B(x_\alpha^k, \delta^k))$$

by the doubling property.

On the other hand, we also know that  $s_\beta^k(x) = 1$  for  $x \in B(x_\beta^k, \frac{1}{8A_0^3} \delta^k)$  and all other  $s_\alpha^k$ ,  $\alpha \neq \beta$ , vanish on this set. Hence

$$\left| \sum_\alpha \lambda_\alpha s_\alpha^k(x) \right|^2 = |\lambda_\beta|^2, \quad \forall x \in B(x_\beta^k, \frac{1}{8A_0^3} \delta^k),$$

thus

$$\left| \sum_\alpha \lambda_\alpha s_\alpha^k(x) \right|^2 \geq \sum_\alpha |\lambda_\alpha|^2 1_{B(x_\alpha^k, \frac{1}{8A_0^3} \delta^k)},$$

and therefore

$$\int \left| \sum_\alpha \lambda_\alpha s_\alpha^k(x) \right|^2 d\mu(x) \geq \sum_\alpha |\lambda_\alpha|^2 \mu(B(x_\alpha^k, \frac{1}{8A_0^3} \delta^k)) \geq \frac{1}{C} \sum_\alpha |\lambda_\alpha|^2 \mu(B(x_\alpha^k, \delta^k)),$$

again by the doubling property.

*The union of the  $V_k$ .* If  $f$  is a Hölder- $\eta$ -continuous function with bounded support, the sum  $f_k(x) = \sum_\alpha f(x_\alpha^k) s_\alpha^k(x)$  defines an element of  $V_k$  and using (3.4) we have

$$f(x) - f_k(x) = \sum_\alpha (f(x) - f(x_\alpha^k)) s_\alpha^k(x)$$

so that

$$\|f - f_k\|_\infty \leq \sup_\alpha \sup_{x \in B(x_\alpha^k, 8A_0^6 \delta^k)} |f(x) - f(x_\alpha^k)| \leq C \delta^{k\eta}.$$

Convergence in  $L^2(\mu)$  follows from this and bounded support for  $f - f_k$ .

*The intersection of the  $V_k$ .* In case  $X$  is bounded, it is clear from definition that the sets  $\{x_\alpha^k\}_\alpha$  reduce to one point when  $k$  gets small. In that case,  $s_\alpha^k(x) = 1$  for all  $x \in X$  by (3.4). So we choose  $k_0$  as the largest integer such that this property holds.

If  $X$  is unbounded, then  $\mu(X) = \infty$  as  $\mu$  is doubling. (See remark after Lemma 8.1 for a proof of this known fact.) If  $f \in \bigcap_{k \in \mathbb{Z}} V_k$ , then  $\|f\|_2^2 \sim \sum_\alpha |f(x_\alpha^k)|^2 \mu_\alpha^k$  for all  $k$ . Consider a fixed  $x \in X$ . Then  $f(x) = \sum_\alpha f(x_\alpha^k) s_\alpha^k(x)$ , where the sum is over the boundedly (with respect to  $k$ ) many  $\alpha$  such that  $x_\alpha^k \in B(x, 8A_0^5 \delta^k)$ . Thus  $|f(x)| \leq \sum_\alpha |f(x_\alpha^k)| \lesssim \sum_\alpha \|f\|_2 / \sqrt{\mu_\alpha^k} \lesssim \|f\|_2 / V(x, \delta^k)$ . This tends to zero as  $k \rightarrow -\infty$ , and hence  $f(x) = 0$ . Since the point was arbitrary, we have  $f \equiv 0$ .  $\square$



## 6. BIORTHOGONAL AND ORTHOGONAL SPLINE SYSTEMS

In this section, we use a classical algorithm (cf. [30, Sec. 2.3]) to construct two further bases of the space  $V_k$  spanned by the splines  $\{s_\alpha^k\}_\alpha$ . We use the abbreviations

$$\mu_\alpha^k := V(x_\alpha^k, \delta^k) := \mu(B(x_\alpha^k, \delta^k))$$

for these frequently appearing volumes.

**6.1. Theorem.** *There exist a system of biorthogonal splines  $\{\tilde{s}_\alpha^k\}_\alpha$  in  $V_k$  with*

$$(6.2) \quad \langle s_\alpha^k, \tilde{s}_\beta^k \rangle_{L^2(\mu)} = \delta_{\alpha,\beta},$$

*as well as an orthonormal basis  $\{\phi_\alpha^k\}_\alpha$  of  $V_k$ , which satisfy the following estimates, where  $d(x, y) \leq \delta^k$ :*

$$\begin{aligned} \mu_\alpha^k \cdot |\tilde{s}_\alpha^k(x)| + \sqrt{\mu_\alpha^k} \cdot |\phi_\alpha^k(x)| &\leq C \exp(-\gamma \delta^{-k} d(x_\alpha^k, x)^s) \\ \mu_\alpha^k \cdot |\tilde{s}_\alpha^k(x) - \tilde{s}_\alpha^k(y)| + \sqrt{\mu_\alpha^k} \cdot |\phi_\alpha^k(x) - \phi_\alpha^k(y)| &\leq C \left( \frac{d(x, y)}{\delta^k} \right)^\eta \exp(-\gamma (\delta^{-k} d(x_\alpha^k, x))^s), \end{aligned}$$

*with  $s = (1 + \log_2 A_0)^{-1}$  or  $s = 1$  if  $d$  is a Lipschitz-continuous quasi-distance.*

**6.3. Remark.** From the decay estimates and doubling, it readily follows that  $\tilde{s}_\beta^k \in L^1(\mu)$ . Summing the biorthogonality relation (6.2) over all  $\alpha$  and recalling that  $\sum_\alpha s_\alpha^k(x) \equiv 1$ , we deduce that

$$\int \tilde{s}_\beta^k d\mu = 1.$$

We begin the proof of Theorem 6.1. Fix  $k \in \mathbb{Z}$ . By abuse, we identify  $\mathcal{X}^k = \{x_\alpha^k\}_\alpha$  and the set of indices  $\alpha$  corresponding to points  $x_\alpha^k$ . The previous section says that there is a linear, bounded, injective map  $U_k: \ell^2(\mathcal{X}^k) \rightarrow L^2(\mu)$  with closed range, defined by

$$U_k \lambda = \sum_\alpha \frac{\lambda_\alpha}{\sqrt{\mu_\alpha^k}} s_\alpha^k$$

for  $\lambda = \{\lambda_\alpha\}_{\alpha \in \mathcal{X}^k}$ . Here,  $\ell^2(\mathcal{X}^k)$  is equipped with counting measure. Let  $V_k$  denote its range. By the properties of the splines, the sum defining each element of  $V_k$  converges locally uniformly and defines a Hölder- $\eta$ -continuous function. The inverse of  $U_k$  can be computed using (3.6) by  $(U_k^{-1}f)_\alpha = f(x_\alpha^k) \sqrt{\mu_\alpha^k}$ .

Let  $\delta_\alpha^k$  be the canonical orthonormal basis element in  $\ell^2(\mathcal{X}^k)$ . By construction  $U_k \delta_\alpha^k = \frac{s_\alpha^k}{\sqrt{\mu_\alpha^k}}$ . Since  $U_k$  is an isomorphism, this means that the splines  $s_\alpha^k$  form an unconditional basis of  $V_k$ . To find a biorthogonal system  $\tilde{s}_\alpha^k$  in  $V_k$ , that is a system such that

$$\langle s_\alpha^k, \tilde{s}_\beta^k \rangle_{L^2(\mu)} = \delta_{\alpha,\beta},$$

we observe that if  $f = U_k \lambda$ ,  $f' = U_k \lambda'$  then

$$\langle f, f' \rangle_{L^2(\mu)} = \langle M_k \lambda, \lambda' \rangle_{\ell^2(\mathcal{X}^k)}$$

with  $M_k$  being the infinite matrix with entries

$$M_k(\alpha, \beta) = \frac{\langle s_\alpha^k, s_\beta^k \rangle_{L^2(\mu)}}{\sqrt{\mu_\alpha^k \mu_\beta^k}}.$$

Another reformulation of the isomorphism property of  $U_k$  is that  $M_k$  is bounded and invertible on  $\ell^2(\mathcal{X}^k)$ . It is also positive self-adjoint. So the biorthogonal system is uniquely defined by

$$\sqrt{\mu_\alpha^k} \tilde{s}_\alpha^k = U_k M_k^{-1} \delta_\alpha^k.$$

If one wants an orthonormal basis of  $V_k \subset L^2(\mu)$ , one defines instead

$$\phi_\alpha^k = U_k M_k^{-1/2} \delta_\alpha^k.$$

In other words,

$$\begin{aligned}\sqrt{\mu_\alpha^k} \tilde{s}_\alpha^k(x) &= \sum_{\beta \in \mathcal{X}^k} M_k^{-1}(\alpha, \beta) \frac{s_\beta^k(x)}{\sqrt{\mu_\beta^k}}, \\ \phi_\alpha^k(x) &= \sum_{\beta \in \mathcal{X}^k} M_k^{-1/2}(\alpha, \beta) \frac{s_\beta^k(x)}{\sqrt{\mu_\beta^k}}.\end{aligned}$$

We now establish decay estimate of the coefficients  $M_k^{-1}(\alpha, \beta)$  and  $M_k^{-1/2}(\alpha, \beta)$ .

**6.4. Proposition.** *There exist constants  $\gamma > 0$  and  $C < \infty$ , independent of  $k$  such that the entries of  $M_k^{-1}$  and  $M_k^{-1/2}$  have upper bounds*

$$C \exp(-\gamma(\delta^{-k} d(\alpha, \beta))^s),$$

with  $s = (1 + \log_2 A_0)^{-1}$  or  $s = 1$  if  $d$  is Lipschitz-continuous.

If we introduce the induced normalized quasi-distance on  $\mathcal{X}^k \times \mathcal{X}^k$ ,  $d_k(\alpha, \beta) := \frac{d(x_\alpha^k, x_\beta^k)}{\delta^k}$ , we have to prove uniform estimates on the entries of  $M_k^{-1}$  or  $M_k^{-1/2}$  in terms of  $d_k$ . Note that for this quasi-distance  $M_k$  is a band-matrix, more precisely  $M_k(\alpha, \beta) = 0$  if  $d_k(\alpha, \beta) \geq 16A_0^6$  by (3.3), hence it has exponential decay.

**6.5. Lemma.** *Let  $\Xi$  be a 1-separated set in a quasi-metric space  $(X, d)$  with quasi-triangle constant  $A_0$  having the geometric doubling property with constant  $N$ . Then for all  $\varepsilon > 0$ , there exists  $c(\varepsilon, A_0, N) < \infty$  such that*

$$\sup_{\alpha \in X} \exp(\varepsilon d(\alpha, \Xi)/A_0) \sum_{\beta \in \Xi} \exp(-\varepsilon d(\alpha, \beta)) \leq c(\varepsilon, A_0, N).$$

Moreover, for any  $0 < c' < c/A_0$

$$\sum_{\beta \in \Xi} \exp(-cd(\alpha, \beta)) \exp(-cd(\beta, \gamma)) \leq C \exp(-c'd(\alpha, \gamma))$$

where  $C$  does not depend on  $\alpha, \gamma \in X$ .

*Proof.* Fix a point  $\alpha \in X$ . Pick any point  $\beta' \in \Xi$ . We have

$$\sum_{\beta \in \Xi} \exp(-\varepsilon d(\alpha, \beta)) \leq \exp(-\varepsilon d(\alpha, \beta')/A_0) \sum_{\beta \in \Xi} \exp(-\varepsilon d(\beta', \beta)/A_0).$$

The number of points in  $\Xi$  at quasi-distance at most  $2^j$  from  $\beta'$  is bounded by  $C^{j+1}$  from some  $C$  depending only on  $A_0$  and  $N$  but not  $\beta'$ . Thus

$$\sum_{\beta \in \Xi} \exp(-\varepsilon d(\beta', \beta)/A_0) \leq C + \sum_{j \geq 1} \sum_{\beta: d(\beta', \beta) \sim 2^j} \exp(-\varepsilon d(\beta', \beta)/A_0) \leq \sum_{j \geq 0} C^{j+1} \cdot \exp(-\varepsilon 2^j/A_0).$$

Taking the infimum over all  $\beta' \in \Xi$  proves the desired inequality.

The result for the matrix product coefficients follows analogously and we skip details.  $\square$

**6.6. Lemma.** *Let  $\Xi$  be a 1-separated set in a quasi-metric space  $(X, d)$  with quasi-triangle constant  $A_0$  having the geometric doubling property with constant  $N$ . Consider a matrix  $M = (M(\alpha, \beta))$  indexed by  $\Xi \times \Xi$  such that there exists  $c > 0$  for which*

$$C = \sup_{(\alpha, \beta)} \exp(cd(\alpha, \beta)) |M(\alpha, \beta)| < \infty.$$

Then  $M$  is bounded on  $\ell^2(\Xi)$ . If  $M$  is invertible, then there exists  $c' > 0$  such that

$$\sup_{(\alpha, \beta)} \exp(c'(d(\alpha, \beta))^s) |M^{-1}(\alpha, \beta)| < \infty$$

with  $s = (1 + \log_2 A_0)^{-1}$  or  $s = 1$  if  $d$  is a Lipschitz-continuous quasi-distance. If, in addition,  $M$  is positive self-adjoint, then the same conclusion holds for  $M^{-1/2}$ .

6.7. *Remark.* (i) Note that the exponent  $s = 1$  for a usual distance is recovered as a special case in two ways, either by setting  $A_0 = 1$  or using the Lipschitz-continuity.

(ii) The exponent  $s$  is in general optimal in this result. Namely, consider the band matrix  $M$  indexed by  $\mathbb{Z}$  with  $M(i, i) = 1$  and  $M(i, i+1) = -\lambda \in (-1, 0)$  for all  $i \in \mathbb{Z}$ , and all other entries equal to zero. Then  $M^{-1}(i, j) = \lambda^{i-j}$  if  $j \geq i$  and zero otherwise. Now equip  $\mathbb{Z}$  with the quasi-distance  $d(i, j) = |i - j|^r$ , where  $r \geq 1$ . Its quasi-triangle constant is  $A_0 = 2^{r-1}$ , and hence the exponent given by the lemma is  $s = (1 + \log_2 A_0)^{-1} = (1 + r - 1)^{-1} = 1/r$ . This gives precisely the correct decay of the matrix with respect to  $d$ , as

$$\lambda^{i-j} = \exp(-c \cdot |i - j|) = \exp(-c \cdot d(i, j)^{1/r}), \quad c = \log \lambda^{-1}.$$

*Proof.* If  $d$  is a genuine distance or a Lipschitz continuous quasi-distance, then this follows from Theorem 5 in [24] and the remark that follows it, which extends an earlier result in [10] for band-limited matrices (but we shall need the full result here), with  $s = 1$ . However, as said not all quasi-distances are Lipschitz- or even Hölder-continuous and we provide an argument in full generality, which also recovers the mentioned special cases.

We begin with the following observation. For  $n \geq 1$ , let  $\kappa_n$  be the best constant in the inequality

$$d(\alpha_0, \alpha_n) \leq \kappa_n(d(\alpha_0, \alpha_1) + d(\alpha_1, \alpha_2) + \dots + d(\alpha_{n-1}, \alpha_n))$$

for every chain  $(\alpha_0, \alpha_1, \dots, \alpha_n)$  of  $n + 1$  elements (not necessarily distinct) of  $\Xi$ . It is clear that  $(\kappa_n)$  is non-decreasing and  $\kappa_1 = 1$ ,  $\kappa_2 \leq A_0$ . Moreover, using  $d(\alpha_0, \alpha_{m+n}) \leq A_0(d(\alpha_0, \alpha_m) + d(\alpha_m, \alpha_{m+n}))$ , it follows that  $\kappa_{m+n} \leq A_0 \cdot \max\{\kappa_m, \kappa_n\}$ . Thus  $\kappa_{2n} \leq A_0 \kappa_n$  and therefore,  $\kappa_{2^j} \leq A_0^j$ . We conclude that  $\kappa_n \leq A_0^{1+\log_2 n} = A_0 n^{\log_2 A_0}$ . Note also that if  $d$  is  $L$ -Lipshitz, then  $d(\alpha_0, \alpha_n) \leq d(\alpha_0, \alpha_{n-1}) + Ld(\alpha_{n-1}, \alpha_n)$ , which by iteration gives  $\kappa_n \leq L$  for all  $n$ .

Now assume that  $M$  is positive self-adjoint and invertible. In this case, one can write  $M = h(I - A)$  with  $h = (\|M\| + \|M^{-1}\|)/2$  a positive real number and  $A$  a matrix with norm  $r = (\|M\| - \|M^{-1}\|)/(\|M\| + \|M^{-1}\|) < 1$ . Moreover, the coefficients of  $A$  have the same decay as those of  $M$ . Without loss of generality, we normalize  $h = 1$ . Develop  $(I - A)^{-1}$  in the Neumann series  $\sum A^n$  and estimate the coefficients  $A^n(\alpha, \beta)$ ,  $n \geq 1$ ,  $\alpha \neq \beta$ , in two ways. First  $|A^n(\alpha, \beta)| \leq r^n$ . Second, we have

$$\begin{aligned} |A^n(\alpha, \beta)| &\leq \sum_{(\alpha_1, \dots, \alpha_{n-1}) \in \Xi^{n-1}} C^n \exp(-c(d(\alpha, \alpha_1) + d(\alpha_1, \alpha_2) + \dots + d(\alpha_{n-1}, \beta))) \\ &\leq C^n \exp(-\frac{c}{2\kappa_n} d(\alpha, \beta)) \sum_{(\alpha_1, \dots, \alpha_{n-1}) \in \Xi^{n-1}} \exp(-\frac{c}{2}(d(\alpha_1, \alpha_2) + \dots + d(\alpha_{n-1}, \beta))) \\ &\leq \tilde{C}^n \exp(-\frac{c}{2\kappa_n} d(\alpha, \beta)), \end{aligned}$$

where we applied Lemma 6.5  $n - 1$  times with  $\varepsilon = c/2$ , and  $\tilde{C} = C \cdot c(\varepsilon, A_0, N)$ , in the notation of that lemma. As  $\kappa_n$  is non decreasing, we have for any integer  $n_0$  using the second estimate for  $0 \leq n \leq n_0$  and the first for  $n > n_0$ ,

$$|M^{-1}(\alpha, \beta)| \leq (n_0 + 1)\tilde{C}^{n_0} \exp(-\frac{c}{2\kappa_{n_0}} d(\alpha, \beta)) + r^{n_0+1}(1 - r)^{-1}$$

and  $(n_0 + 1)\tilde{C}^{n_0} \leq D^{n_0}$  for some large constant  $D > 0$ . Choosing  $n_0$  as the first integer such that the first term dominates, we see that  $d(\alpha, \beta) \approx n_0 \cdot \kappa_{n_0} \lesssim n_0^{1/s}$ , where  $s = 1/(1 + \log_2 A_0)$  (or  $s = 1$  if  $d$  is Lipschitz-continuous, recalling that  $\kappa_{n_0} \leq L$  in this case). Hence, for some constant  $c' > 0$ ,

$$|M^{-1}(\alpha, \beta)| \lesssim r^{n_0} \lesssim \exp(-c' d(\alpha, \beta)^s).$$

For  $M^{-1/2}$ , we use the power series  $(I - A)^{-1/2} = \sum c_n A^n$ . As  $0 \leq c_n \lesssim n^{1/2}$ , the argument is the same. Finally, if  $M$  is not positive self-adjoint, then we use  $M^{-1} = M^*(MM^*)^{-1}$  and the remark that on a 1-separated set  $d \geq d^s$ .  $\square$

*Proof of Proposition 6.4 and Theorem 6.1.* Lemma 6.6 gives us the desired decay for  $M_k^{-1}$  and  $M_k^{-1/2}$ , as claimed in Proposition 6.4, with uniform control of the constant with respect to  $k$ . The

decay and regularity of the  $\tilde{s}_\alpha^k$  and  $\phi_\alpha^k$ , as asserted by Theorem 6.1, then follow from the support and regularity of the  $s_\alpha^k$  and the decay of the matrix coefficients. We skip details.  $\square$

6.8. *Remark.* Let  $0 < s \leq 1$ . If  $d$  is a quasi-distance with quasi-triangle constant  $A_0$ , then  $d^s$  is a quasi-distance with quasi-triangle constant  $A_0^s$ . If  $d$  is geometrically doubling with constant  $N$ , then so is  $d^s$  with constant at most  $N^{\lceil 1/s \rceil}$ , where  $\lceil \cdot \rceil$  is the ‘rounding-up’ to the next integer. If  $\mu$  is a doubling measure with respect to the balls for  $d$  with constant  $c_\mu$ , then it is with respect to the balls for  $d^s$  with constant at most  $c_\mu^{\lceil 1/s \rceil}$ . Thus both previous lemmata apply to  $d^s$  with constants that depend only on the original constants and  $s$ . We shall use this remark later.

## 7. SPLINE WAVELETS

We are now prepared for the construction of an orthonormal basis of  $L^2(\mu)$ , consisting of wavelets  $\psi_\alpha^k$  with similar decay and regularity properties as with the spline systems. We follow an algorithm from Meyer [29].

Fix  $k \in \mathbb{Z}$ ,  $k \geq k_0$ , where  $k_0$  is as in Theorem 5.1. Recall the operator

$$U_{k+1}^{-1} : f \mapsto \{f(x_\beta^{k+1})\sqrt{\mu_\beta^{k+1}}\}_\beta$$

is an isomorphism from  $V_{k+1}$  onto  $\ell^2(\mathcal{X}^{k+1})$ . Denote by  $Y_k$  the inverse image of the subspace of  $\ell^2(\mathcal{X}^{k+1})$  sequences vanishing on  $\mathcal{X}^k$ ; this subspace is naturally identified with  $\ell^2(\mathcal{Y}^k)$ , where we recall the notation  $\mathcal{Y}^k := \mathcal{X}^{k+1} \setminus \mathcal{X}^k$ . Clearly,  $V_k \oplus Y_k = V_{k+1}$  topologically. Consider the orthogonal (in  $L^2(\mu)$ ) complement  $W_k$  of  $V_k$  in  $V_{k+1}$ . Then the restriction to  $Y_k$  of the orthogonal projection  $Q_k$  onto  $W_k$  is an isomorphism onto  $W_k$ . Now, identifying with  $\mathcal{Y}^k$  the set of indices  $\beta$  corresponding to  $x_\beta^{k+1} \in \mathcal{X}^{k+1} \setminus \mathcal{X}^k = \mathcal{Y}^k$ , the collection  $\{s_\beta^{k+1}\}_{\beta \in \mathcal{Y}^k}$  forms an unconditional basis of  $Y_k$  and its image under  $Q_k$  is an unconditional basis of  $W_k$ . A representation of  $Q_k f$  when  $f \in V_{k+1}$  is

$$(7.1) \quad Q_k f = f - \sum_{\alpha \in \mathcal{X}^k} \langle f, \tilde{s}_\alpha^k \rangle_{L^2(\mu)} s_\alpha^k = f - \sum_{\alpha \in \mathcal{X}^k} \langle f, s_\alpha^k \rangle_{L^2(\mu)} \tilde{s}_\alpha^k,$$

because the sum is the orthogonal projection of  $f$  onto  $V_k$ . Hence, the pre-wavelets

$$\tilde{\psi}_\beta^k(x) := Q_k s_\beta^{k+1}(x)$$

have the  $L^\infty$ -normalized exponential decay  $C \exp(-\gamma(\delta^{-k} d(y_\beta^k, x))^s)$ , where  $y_\beta^k := x_\beta^{k+1} \in \mathcal{Y}^k$ . (Remark that they are normalized as the splines  $s_\beta^{k+1}$ .) Finally, one can orthonormalize them in  $L^2(\mu)$  following the procedure of Section 6 applied to the positive self-adjoint matrix

$$\tilde{M}(\alpha, \beta) := \frac{\langle \tilde{\psi}_\alpha^k, \tilde{\psi}_\beta^k \rangle}{\sqrt{\mu_\alpha^{k+1} \mu_\beta^{k+1}}}$$

indexed by  $\mathcal{Y}^k \times \mathcal{Y}^k$ ; that is, we define

$$\psi_\alpha^k(x) := \sum_{\beta \in \mathcal{Y}^k} \tilde{M}^{-1/2}(\alpha, \beta) \frac{\tilde{\psi}_\beta^k(x)}{\sqrt{\mu_\beta^{k+1}}}.$$

Note that with the notation  $y_\beta^k = x_\beta^{k+1}$ , we have  $\mu_\beta^{k+1} = \mu(B(y_\beta^k, \delta^{k+1})) \sim \mu(B(y_\beta^k, \delta^k))$ . The point-set  $\mathcal{Y}^k$  is a 1-separated set for  $d_k^s(x, y) := (\delta^{-k-1} d(x, y))^s$  and the matrix  $\tilde{M}$  has the exponential decay, as required in Lemma 6.6, with respect to  $d_k^s$ . By Lemma 6.6 and Remark 6.8,  $\tilde{M}^{-1/2}$  has decay

$$|\tilde{M}^{-1/2}(\alpha, \beta)| \lesssim \exp(-\gamma' d_k^s(x_\alpha^k, x_\beta^k)^{\tilde{s}}) = \exp(-\gamma(\delta^{-k} d(x_\alpha^k, x_\beta^k))^{s\tilde{s}}),$$

where  $\tilde{s} = (1 + \log_2 A_0^s)^{-1} = (1 + (1 + \log_2 A_0)^{-1} \log_2 A_0)^{-1} = (1 + \log_2 A_0)(1 + 2 \log_2 A_0)^{-1}$  or  $\tilde{s} = 1$  if  $d$  is Lipschitz-continuous.

This yields an orthonormal basis  $\psi_\beta^k(x)$ ,  $y_\beta^k \in \mathcal{Y}^k$ , of  $W_k$  having the  $L^2$  normalized exponential decay

$$|\psi_\beta^k(x)| \leq \frac{C}{\sqrt{\mu(B(y_\beta^k, \delta^k))}} \exp(-\gamma(\delta^{-k}d(y_\beta^k, x))^a), \quad a = s\tilde{s}.$$

Gathering the construction for all  $k$  (and adding constants if  $X$  is bounded), one obtains an orthonormal basis of  $L^2(\mu)$  made of spline wavelets.

**7.2. Theorem.** *Let  $(X, d, \mu)$  be any space of homogeneous type with quasi-triangle constant  $A_0$ , and  $a := (1 + 2\log_2 A_0)^{-1}$  or  $a := 1$  if  $d$  is Lipschitz-continuous. There exists an orthonormal basis  $\psi_\beta^k$ ,  $k \in \mathbb{Z}$  (and  $k \geq k_0$  if  $X$  is bounded),  $y_\beta^k \in \mathcal{Y}^k$ , of  $L^2(\mu)$  (or the orthogonal space to constants if  $X$  is bounded) having exponential decay*

$$|\psi_\beta^k(x)| \leq \frac{C}{\sqrt{\mu(B(y_\beta^k, \delta^k))}} \exp(-\gamma(\delta^{-k}d(y_\beta^k, x))^a),$$

*Hölder-regularity*

$$|\psi_\beta^k(x) - \psi_\beta^k(y)| \leq \frac{C}{\sqrt{\mu(B(y_\beta^k, \delta^k))}} \left(\frac{d(x, y)}{\delta^k}\right)^\eta \exp(-\gamma(\delta^{-k}d(y_\beta^k, x))^a), \quad d(x, y) \leq \delta^k,$$

*and vanishing mean*

$$\int_X \psi_\beta^k(x) d\mu(x) = 0, \quad k \in \mathbb{Z}, k \geq k_0, y_\beta^k \in \mathcal{Y}^k.$$

*Proof.* It remains to see vanishing mean and regularity.

Using  $\int \tilde{s}_\alpha^k d\mu = 1$  and  $\sum_\alpha s_\alpha^k \equiv 1$ , we deduce from (7.1) for  $f \in L^1(\mu)$  that

$$\int Q_k f d\mu = \int f d\mu - \sum_\alpha \langle f, s_\alpha^k \rangle \int \tilde{s}_\alpha^k d\mu = \int f d\mu - \sum_\alpha \int f s_\alpha^k d\mu = \int f d\mu - \int f d\mu = 0.$$

Since the pre-wavelets  $\tilde{\psi}_\alpha^k$  lie in the range of  $Q_k$  by definition, we have  $\int \tilde{\psi}_\alpha^k d\mu = 0$ , and the same result for the wavelets follows from the convergent series representation of  $\psi_\alpha^k$  in terms of the  $\tilde{\psi}_\beta^k$ .

As for the regularity, recall the smoothness of the splines:  $|s_\alpha^k(x) - s_\alpha^k(y)| \leq C(d(x, y)/\delta^k)^\eta$ . This implies for the pre-wavelets the estimate

$$\begin{aligned} |\tilde{\psi}_\beta^k(x) - \tilde{\psi}_\beta^k(y)| &\leq |s_\beta^{k+1}(x) - s_\beta^{k+1}(y)| + \sum_\alpha |\langle s_\beta^{k+1}, \tilde{s}_\alpha^k \rangle| |s_\alpha^k(x) - s_\alpha^k(y)| \\ &\leq C\left(\frac{d(x, y)}{\delta^k}\right)^\eta 1_{B(y_\beta^k, c\delta^k)}(x) + \sum_\alpha C \exp(-\gamma(d(y_\beta^k, x_\alpha^k)/\delta^k)^s) \left(\frac{d(x, y)}{\delta^k}\right)^\eta 1_{B(x_\alpha^k, C\delta^k)}(x) \\ &\leq C\left(\frac{d(x, y)}{\delta^k}\right)^\eta \exp(-\gamma(d(x, y_\beta^k)/\delta^k)^s). \end{aligned}$$

For the wavelets, finally, we have

$$\begin{aligned} |\psi_\alpha^k(x) - \psi_\alpha^k(y)| &\leq C \sum_\beta \frac{\exp(-\gamma(d(y_\alpha^k, y_\beta^k)/\delta^{k+1})^a)}{\sqrt{\mu(B(y_\beta^k, \delta^k))}} |\tilde{\psi}_\beta^k(x) - \tilde{\psi}_\beta^k(y)| \\ &\leq C \sum_\beta \frac{\exp(-\gamma(d(y_\alpha^k, y_\beta^k)/\delta^{k+1})^a)}{\sqrt{\mu(B(y_\beta^k, \delta^k))}} \left(\frac{d(x, y)}{\delta^k}\right)^\eta \exp(-\gamma(d(x, y_\beta^k)/\delta^k)^s) \\ &\leq \frac{C}{\sqrt{\mu(B(y_\alpha^k, \delta^k))}} \left(\frac{d(x, y)}{\delta^k}\right)^\eta \exp(-\gamma(d(x, y_\alpha^k)/\delta^k)^a). \end{aligned}$$

Note that the value of  $\gamma > 0$  changes from line to line in these computations. Also we used that  $a \leq s$  and a variant of Lemma 6.5 for the quasi-distance  $d(x, y)/\delta^k$  on  $X$ .  $\square$

**7.3. Remark.** The construction and also the next sections suggest that the label  $k + 1$  would be more appropriate than  $k$  for  $W_k$ , the wavelets  $\psi_\alpha^k$  and their scale  $\delta^k$ , because it keeps closer to the definition of the point sets  $\mathcal{Y}^k$ , a subset of  $\mathcal{X}^{k+1}$ , which will take an important role. We have kept the wavelet community notation as in [30].

## 8. TECHNICAL ESTIMATES RELATED TO VANISHING ANNULI

We break the development of the wavelet theory with this technical section, which will provide us with useful estimates to streamline the subsequent presentation. A basic difficulty related to general spaces of homogeneous type, as opposed to those with the reverse doubling property, is the possible existence of arbitrarily large empty annuli  $B(x, R) \setminus B(x, r) = \emptyset$ . This leads to a certain dichotomy: locally, we either have the reverse doubling estimate, or the vanishing of a certain annulus, both of which provide certain control, which we need to exploit in different ways. This is quantified in the following. (The next result is well-known, cf. [14, Remark 1.2], but we include it here for completeness, since we need it to derive some consequences which appear to be new.)

**8.1. Lemma.** *For every  $x \in X$  and  $R > r > 0$ , at least one of the following alternatives holds:*

$$\mu(B(x, R)) \geq (1 + \varepsilon)\mu(B(x, r)) \quad \text{or} \quad B(x, \frac{1}{2A_0}R) \setminus B(x, 2A_0r) = \emptyset,$$

where  $\varepsilon := 1/C_\mu(3A_0^2)$ , and  $C_\mu(t)$  is the smallest constant such that  $\mu(tB) \leq C_\mu(t)\mu(B)$  for all balls  $B \subseteq X$ .

In particular, as  $R \rightarrow \infty$ , we observe the following (again well-known fact): If there is a ball  $B(x, r)$  such that  $\mu(X) < (1 + C_\mu(3A_0^2)^{-1})\mu(B(x, r))$ , then  $X = B(x, 2A_0r)$ . If  $\mu(X) < \infty$ , such a ball always exists, and hence  $\text{diam}(X) < \infty$ .

*Proof.* Suppose that the annulus is nonempty, and let  $y \in B(x, \frac{1}{2A_0}R) \setminus B(x, 2A_0r)$ . Let  $\rho := \frac{1}{2A_0}d(x, y) \geq r$ . We claim that

$$B(y, \rho) \subseteq B(x, R) \setminus B(x, r).$$

Indeed, if  $z \in B(y, \rho)$ , then

$$d(z, x) \leq A_0d(z, y) + A_0d(y, x) < A_0\rho + A_0d(x, y) \leq R,$$

while

$$d(z, x) \geq \frac{1}{A_0}d(y, x) - d(y, z) > \frac{1}{A_0}d(x, y) - \rho = \frac{1}{2A_0}d(x, y) \geq r.$$

We also claim that

$$B(x, r) \subset B(y, 3A_0^2\rho).$$

Indeed, if  $w \in B(x, r)$ , then

$$d(w, y) \leq A_0d(w, x) + A_0d(x, y) < A_0r + 2A_0^2\rho \leq 3A_0^2\rho.$$

Now it follows that

$$\mu(B(x, R)) - \mu(B(x, r)) \geq \mu(B(y, \rho)) \geq \frac{1}{C_\mu(3A_0^2)}\mu(B(y, 3A_0^2\rho)) \geq \frac{1}{C_\mu(3A_0^2)}\mu(B(x, r)). \quad \square$$

The following lemma relates the mentioned dichotomy to the distribution of the dyadic point sets  $\mathcal{Y}^k$ :

**8.2. Lemma.** *For every  $x \in X$  and  $r > 0$ , there exists a decreasing sequence, finite or infinite, of integers  $\{k_j\}_{j=0}^J$  such that  $r \leq \delta^{k_0} < \delta^{k_1} < \dots$  such that*

$$V(x, \delta^k) \gtrsim (1 + \varepsilon)^j V(x, r) \quad \text{and} \quad d(x, \mathcal{Y}^k) + \delta^k \gtrsim \delta^{k_{j+1}} \quad \text{if } k_j \geq k > k_{j+1},$$

where we interpret  $k_{J+1} := -\infty$  if  $J < \infty$ .



*Proof.* Let  $k(0)$  be the largest integer with  $\delta^{k(0)} \geq r$ , and let  $k(j+1)$  be the largest integer with  $V(x, \delta^{k(j+1)}) \geq (1+\varepsilon)V(x, \delta^{k(j)})$ . (Note that the sequence terminates if and only if  $\mu(X) < \infty$ .) Thus  $V(x, \delta^{k(j+1)+1}) < (1+\varepsilon)V(x, \delta^{k(j)})$ , and hence

$$B(x, \frac{1}{2A_0}\delta^{k(j+1)+1}) = B(x, 2A_0\delta^{k(j)}).$$

For  $k \geq k(j+1)+2$ , the ball on the left contains at least one element of  $\mathcal{X}^k$ . For  $k \leq k(j)-1$ , the ball on the right contains at most one element of  $\mathcal{X}^k$ . Since the balls are equal, for  $k(j+1)+2 \leq k \leq k(j)-1$ , the ball contains exactly one element of  $\mathcal{X}^k$ . So if  $k$  and  $k+1$  are both in this range, i.e., if  $k(j+1)+2 \leq k \leq k(j)-2$ , then the intersections of the ball with  $\mathcal{X}^k$  and  $\mathcal{X}^{k+1}$  coincide; hence there is no point of  $\mathcal{Y}^k = \mathcal{X}^{k+1} \setminus \mathcal{X}^k$  in the ball. Thus  $d(x, \mathcal{Y}^k) \geq \frac{1}{2A_0}\delta \cdot \delta^{k(j+1)}$ . On the other hand, if  $k \in \{k(j+1)-1, k(j+1), k(j+1)+1\}$ , then clearly  $\delta^k \geq \delta \cdot \delta^{k(j+1)}$ , and hence

$$d(x, \mathcal{Y}^k) + \delta^k \geq \frac{\delta}{2A_0}\delta^{k(j+1)} \quad \text{if } k(j+1)-1 \leq k \leq k(j)-2.$$

Also, for  $k$  in the same range, we have

$$V(x, \delta^k) \geq V(x, \delta^{k(j)}) \geq (1+\varepsilon)^j V(x, \delta^{k(0)}) \geq (1+\varepsilon)^j V(x, r).$$

So the claim follows by relabeling  $k_{j+1} := k(j) - 2$  for  $j \in \mathbb{N}$  and  $k_0 := k(0)$ .  $\square$

Sums of the following type appear in connection with the wavelets:

**8.3. Lemma.** *For all  $x \in X$  and  $r, \nu, a > 0$ , we have*

$$\sum_{k: \delta^k \geq r} V(x, \delta^k)^{-\nu} \exp(-\gamma(\delta^{-k}d(x, \mathcal{Y}^k))^a) \lesssim V(x, r)^{-\nu}.$$

*Proof.* Let  $k_j$  be the sequence as provided by Lemma 8.2. Then

$$\begin{aligned} & \sum_{k: \delta^k \geq r} V(x, \delta^k)^{-\nu} \exp(-\gamma(\delta^{-k}d(x, \mathcal{Y}^k))^a) \\ &= \sum_{j=0}^J \sum_{k: k_j \geq k > k_{j+1}} V(x, \delta^k)^{-\nu} \exp(-\gamma(\delta^{-k}d(x, \mathcal{Y}^k))^a) \\ &\lesssim \sum_{j=0}^J \sum_{k: k_j \geq k > k_{j+1}} (1+\varepsilon)^{-j\nu} V(x, r)^{-\nu} \exp(-\gamma\delta^{a(k_{j+1}-k)}) \\ &\leq V(x, r)^{-\nu} \left( \sum_{j=0}^J (1+\varepsilon)^{-j\nu} \right) \left( \sum_{m=1}^{\infty} \exp(-\gamma\delta^{-ma}) \right) \lesssim V(x, r)^{-\nu}. \end{aligned} \quad \square$$

## 9. TECHNICAL ESTIMATES INVOLVING THE WAVELETS

Actually, the estimates here are valid for any family of functions  $\psi_\alpha^k$  which satisfy the same size and regularity estimates as the wavelets. This includes the condition that  $\psi_\alpha^k$  be concentrated around the point  $y_\alpha^k \in \mathcal{Y}^k$ , and the structure of the point sets  $\mathcal{Y}^k$  is important for some of the following estimates.

**9.1. Lemma.** *Let  $a$  as in Theorem 7.2. For a fixed  $k \in \mathbb{Z}$ ,*

$$\sum_{\alpha \in \Lambda_k} |\psi_\alpha^k(x) \psi_\alpha^k(y)| \leq \frac{C}{V(x, \delta^k)} \exp(-\gamma(\delta^{-k}d(x, \mathcal{Y}^k))^a) \exp(-\gamma(\delta^{-k}d(x, y))^a),$$

and, for  $d(x, x') < \frac{1}{2A_0}d(x, y)$ ,

$$\begin{aligned} & \sum_{\alpha \in \Lambda_k} |[\psi_\alpha^k(x) - \psi_\alpha^k(x')] \psi_\alpha^k(y)| \\ & \leq \frac{C}{V(x, \delta^k)} \min \left\{ 1, \left( \frac{d(x, x')}{\delta^k} \right)^\eta \right\} \exp(-\gamma(\delta^{-k}d(x, \mathcal{Y}^k))^a) \exp(-\gamma(\delta^{-k}d(x, y))^a). \end{aligned}$$

Note that if, instead, we have a family of functions  $\varphi_\alpha^k$  corresponding to the points  $x_\alpha^k \in \mathcal{X}^k$  instead of  $y_\alpha^k \in \mathcal{Y}^k$ , we get exactly the same estimate but with  $d(x, \mathcal{X}^k)$  in place of  $d(x, \mathcal{Y}^k)$  in the result. Since  $d(x, \mathcal{X}^k) \leq 2A_0\delta^k$ , the first exponential factor is roughly 1, and may be dropped.

*Proof.* By the doubling condition and the quasi-triangle inequality, we estimate

$$\begin{aligned} |\psi_\alpha^k(x)\psi_\alpha^k(y)| &\leq \frac{C}{\sqrt{\mu_\alpha^{k+1}}} \exp(-\gamma(\delta^{-k}d(x, y_\alpha^k))^a) \frac{C}{\sqrt{\mu_\alpha^{k+1}}} \exp(-\gamma(\delta^{-k}d(y, y_\alpha^k))^a) \\ &\leq \frac{C}{V(x, \delta^k)} \exp(-\gamma(\delta^{-k}d(x, y_\alpha^k))^a) \exp(-\gamma(\delta^{-k}d(x, y))^a), \end{aligned}$$

and the sum over  $\alpha \in \mathcal{Y}^k$  of the second factor is dominated by  $\exp(-\gamma(\delta^{-k}d(x, \mathcal{Y}^k))^a)$ .

If  $\delta^k \geq d(x, x')$ , then

$$\begin{aligned} |[\psi_\alpha^k(x) - \psi_\alpha^k(x')]\psi_\alpha^k(y)| &\leq \frac{C}{\mu_\alpha^{k+1}} \left(\frac{d(x, x')}{\delta^k}\right)^\eta \exp(-\gamma(\delta^{-k}d(x, y_\alpha^k))^a) \exp(-\gamma(\delta^{-k}d(y, y_\alpha^k))^a) \\ &\leq \frac{C}{V(x, \delta^k)} \left(\frac{d(x, x')}{\delta^k}\right)^\eta \exp(-\gamma(\delta^{-k}d(x, y_\alpha^k))^a) \exp(-\gamma(\delta^{-k}d(x, y))^a), \end{aligned}$$

and we may similarly sum over  $\alpha \in \mathcal{Y}^k$ . For  $\delta^k < d(x, x') \leq d(x, y)$ , we just use the quasi-triangle inequality and the first estimate of the lemma to both terms.  $\square$

## 9.2. Lemma.

$$\sum_{k, \alpha} |\psi_\alpha^k(x)\psi_\alpha^k(y)| \leq \frac{C}{V(x, y)}.$$

*Proof.* By Lemma 9.1, we have

$$\begin{aligned} \sum_{k, \alpha} |\psi_\alpha^k(x)\psi_\alpha^k(y)| &\leq \sum_{k: \delta^k \geq d(x, y)} \frac{C}{V(x, \delta^k)} \exp(-\gamma(\delta^{-k}d(x, \mathcal{Y}^k))^a) \\ &\quad + \sum_{k: \delta^k < d(x, y)} \frac{C}{V(x, \delta^k)} \exp(-\gamma(\delta^{-k}d(x, y))^a), \end{aligned}$$

where the first part has the correct bound by Lemma 8.3. For the second part, we have

$$\begin{aligned} &\sum_{k: \delta^k < d(x, y)} \frac{C}{V(x, \delta^k)} \exp(-\gamma(\delta^{-k}d(x, y))^a) \\ &\leq \sum_{k: \delta^k < d(x, y)} \frac{C}{V(x, y)} \left(\frac{d(x, y)}{\delta^k}\right)^M \exp(-\gamma(\delta^{-k}d(x, y))^a) \\ &\leq \frac{C}{V(x, y)} \sum_{m=0}^{\infty} \delta^{-mM} \exp(-\gamma\delta^{-ma}) \leq \frac{C}{V(x, y)}. \end{aligned} \quad \square$$

## 9.3. Lemma.

$$\sum_{k, \alpha} |[\psi_\alpha^k(x) - \psi_\alpha^k(x')]\psi_\alpha^k(y)| \leq \frac{C}{V(x, y)} \left(\frac{d(x, x')}{d(x, y)}\right)^\eta, \quad d(x, x') < \frac{1}{2A_0}d(x, y).$$

*Proof.* By the second estimate of Lemma 9.1,

$$\begin{aligned}
& \sum_{k,\alpha} |[\psi_\alpha^k(x) - \psi_\alpha^k(x')]\psi_\alpha^k(y)| \\
& \leq \sum_{k:\delta^k \geq d(x,y)} \frac{C}{V(x,\delta^k)} \left(\frac{d(x,x')}{d(x,y)}\right)^\eta \exp(-\gamma(\delta^{-k}d(x,\mathscr{Y}^k))^a) \\
& \quad + \sum_{k:d(x,x') \leq \delta^k \leq d(x,y)} \frac{C}{V(x,y)} \left(\frac{d(x,y)}{\delta^k}\right)^{M+\eta} \left(\frac{d(x,x')}{d(x,y)}\right)^\eta \exp(-\gamma(\delta^{-k}d(x,y))^a) \\
& \quad + \sum_{k:\delta^k \leq d(x,x')} \frac{C}{V(x,y)} \left(\frac{d(x,y)}{\delta^k}\right)^M \exp(-\gamma(\delta^{-k}d(x,y))^a).
\end{aligned}$$

The first two parts contain the factor  $(d(x,x')/d(x,y))^\eta$ , and the rest is bounded by  $C/V(x,y)$  according to Lemma 8.3 and  $\sum_{m=0}^\infty \delta^{-m(M+\eta)} \exp(-\gamma\delta^{-ma}) \leq C$ . For the last term we even obtain the bound  $C/V(x,y) \cdot \exp(-\gamma(d(x,y)/d(x,x'))^a) \leq C/V(x,y) \cdot (d(x,x')/d(x,y))^K$  for any  $K$ .  $\square$

## 10. LITTLEWOOD–PALEY DECOMPOSITION AND $L^p$ THEORY

Recall that  $Q_k$  is the orthogonal projector onto  $W_k$  and set also  $P_k$  the orthogonal projector onto  $V_k$ . These operators will provide us with a new regular Littlewood–Paley decomposition for spaces of homogeneous type. The following lemma describes the kernels of these operators:

**10.1. Lemma.** *The kernel  $P_k(x,y)$  of  $P_k$  is symmetric in  $x,y$  and has estimates*

$$\begin{aligned}
|P_k(x,y)| & \leq \frac{C}{\sqrt{\mu(B(x,\delta^k))\mu(B(y,\delta^k))}} \exp(-\gamma(\delta^{-k}d(x,y))^s) \\
|P_k(x,y) - P_k(x,y')| & \leq C \left(\frac{d(y,y')}{\delta^k}\right)^\eta \left( \frac{\exp(-\gamma(\delta^{-k}d(x,y))^s)}{\sqrt{\mu(B(x,\delta^k))\mu(B(y,\delta^k))}} + \frac{\exp(-\gamma(\delta^{-k}d(x,y'))^s)}{\sqrt{\mu(B(x,\delta^k))\mu(B(y',\delta^k))}} \right)
\end{aligned}$$

for some  $C, \gamma$  and all  $x, y, y' \in X$  and  $k \in \mathbb{Z}$  (with  $k \geq k_0$  if  $X$  is bounded). Moreover

$$\int_X P_k(x,y) d\mu(x) = 1.$$

The kernel  $Q_k(x,y)$  of  $Q_k$  is symmetric in  $x,y$  and has similar estimates with  $s$  changed to  $a$ , the additional exponential factor

$$\exp(-\gamma(\delta^{-k}d(x,\mathscr{Y}^k))^a),$$

and the cancellation condition

$$\int_X Q_k(x,y) d\mu(x) = 0.$$

Note that there are many ways to express the denominator factors in the kernels up to changing the constants  $C, \gamma$  in the numerator factors especially thanks to the exponential decay. In particular, one possible expression shows that the system of operators  $P_k$  is an ‘Approximation of The Identity’ in the sense of [15, Definition 2.2] (this part of that paper does not use the Reverse Doubling property), and the properties listed there hold.

**10.2. Theorem.** *The spline-wavelet representation yields a decomposition of Littlewood–Paley type with Hölder-continuous kernels.*

*Proof.* If  $X$  is unbounded and  $f \in L^2(X)$  the converging series

$$f = \sum_{k=-\infty}^{\infty} Q_k f = \sum_{k=-\infty}^{\infty} Q_k^2 f$$

is an homogeneous Littlewood-Paley decomposition. Also one can truncate at any level since  $P_{k+1} = P_k + Q_k$  and write for any  $\ell$

$$f = P_\ell^2 f + \sum_{k=\ell}^{\infty} Q_k^2 f$$

which is an inhomogeneous Littlewood-Paley decomposition. This decomposition is the one used if  $X$  is bounded with  $\ell = k_0$ . In that case, the term  $P_\ell^2 f$  is a constant.  $\square$

Observe that the sum  $\sum_k Q_k f$  can also be rewritten as a “discrete” Littlewood-Paley decomposition  $\sum_{k,\alpha} \langle f, \psi_\alpha^k \rangle \psi_\alpha^k$ . From there, one can look at convergence for  $f$  in various topological spaces, and develop the theory of function spaces. We restrict ourselves to  $L^p$  spaces (and BMO in the next section) and leave further developments to the interested reader.

The estimates in Section 9 immediately give the following result:

**10.3. Proposition.** *Let  $c_\alpha^k$  be arbitrary complex coefficients bounded in absolute value by one. Then the series*

$$K(x, y) = \sum_{k \in \mathbb{Z}} \sum_{\alpha \in \Lambda_k} c_\alpha^k \psi_\alpha^k(x) \psi_\alpha^k(y)$$

*converges absolutely for  $x \neq y$  and satisfies*

$$|K(x, y)| \leq \frac{C}{V(x, y)}, \quad |K(x, y) - K(x', y)| \leq \frac{C}{V(x, y)} \left( \frac{d(x, x')}{d(x, y)} \right)^\eta, \quad d(x, x') \leq \frac{1}{2A_0} d(x, y),$$

*with a similar Hölder-regularity estimate in the second variable.*

**10.4. Corollary.** *The spline wavelets form an unconditional basis of  $L^p(\mu)$  spaces when  $1 < p < \infty$ .*

*Proof.* Completeness follows using the convergence properties of the  $P_k$  in  $L^p(\mu)$ . It remains to show that operators  $T_c$  given  $T_c(\psi_\alpha^k) = c_\alpha^k \psi_\alpha^k$  are uniformly bounded on all  $L^p(\mu)$  spaces whenever  $c = (c_\alpha^k)$  is a sequence of complex numbers in the unit ball of  $\mathbb{C}$ . These operators are contractions in  $L^2(\mu)$ . Using the regularity of their kernels as proved in Proposition 10.3 and the Calderón–Zygmund theorem,  $T_c$  has weak type (1,1), with uniform bound with respect to  $c$ . Since  $T_c^* = T_{\bar{c}}$ , the same applies to  $T_c^*$ . We conclude by interpolation.  $\square$

## 11. BMO THEORY

Recall that the space  $\text{BMO}(\mu)$  of functions of **bounded mean oscillation** consists of those  $b \in L_{\text{loc}}^1(\mu)$  with

$$\|b\|_{\text{BMO}(\mu)} := \sup_B \inf_c \frac{1}{\mu(B)} \int_B |b - c| d\mu < \infty,$$

where the infimum is almost realized by  $c = b_B := \mu(B)^{-1} \int_B b d\mu$ . We do not incorporate the norm of constants as a part of the BMO norm even if  $X$  is bounded, so that our BMO is a Banach space modulo constants in both bounded and unbounded cases. These averages satisfy the following useful estimate:

**11.1. Lemma.** *Let  $B_i = B(x_i, r_i)$ ,  $i = 1, 2$ , be two balls in  $X$ . Then*

$$|b_{B_1} - b_{B_2}| \lesssim \|b\|_{\text{BMO}(\mu)} \left( 1 + \log \frac{r_1 + r_2 + d(x_1, x_2)}{\min\{r_1, r_2\}} \right).$$

*Proof.* Suppose first that  $r_1 \leq r_2$  and  $B_1 \subset B_2$ . Let  $B^0 := B_1$  and  $B^{i+1}$  be the smallest  $2^j B^i$  such that  $\mu(B^{i+1}) > C_\mu \mu(B^i)$ ; hence also  $\mu(B^{i+1}) \leq C_\mu^2 \mu(B^i)$ . It is easy to check that  $|b_{B^i} - b_{B^{i+1}}| \lesssim \|b\|_{\text{BMO}(\mu)}$ . Let  $B^k$  be the first  $B^i$  with radius bigger than that of  $B_2$ . Then we also have  $|b_{B^k} - b_{B_2}| \lesssim \|b\|_{\text{BMO}(\mu)}$ , and  $k \lesssim 1 + \log r_2/r_1$ .

In the general case, we may choose an auxiliary ball  $B_3$  of radius  $r_3 \sim r_1 + r_2 + d(x_1, x_2)$  which contains both  $B_1$  and  $B_2$ , and apply the earlier consideration to

$$|b_{B_1} - b_{B_2}| \leq |b_{B_1} - b_{B_3}| + |b_{B_3} - b_{B_2}|. \quad \square$$

**11.2. Corollary.** *Let  $f$  be a function with exponential decay  $|f(x)| \leq C \exp(-c \cdot d(x, x_0)^a)$  with  $C, c, a > 0$ , and  $b \in \text{BMO}(\mu)$ . Then the product  $f \cdot b$  is integrable over  $X$ .*

*Proof.* Let  $B_n := B(x_0, n)$ . Clearly  $f$ , and hence  $f \cdot b_{B_1}$  is integrable, so we consider  $f \cdot (b - b_{B_1})$ . By Lemma 11.1,

$$\int_{B_n} |b - b_{B_1}| d\mu \leq \int_{B_n} |b - b_{B_n}| d\mu + \mu(B_n) |b_{B_n} - b_{B_1}| \lesssim \mu(B_n) (1 + \log n) \|b\|_{\text{BMO}(\mu)},$$

and  $\mu(B_n) \lesssim n^M \mu(B_0)$  by doubling. Hence  $\|1_{B_n}(b - b_{B_0})\|_{L^1(\mu)}$  grows at most polynomially in  $n$ , whereas  $\|1_{B_n \setminus B_{n-1}} f\|_{L^\infty(\mu)}$  decays exponentially in  $n$ . Thus  $f \cdot (b - b_{B_0}) \in L^1(\mu)$ .  $\square$

This implies in particular that the wavelet coefficients  $(b, \psi_\alpha^k) = \int_X b \cdot \psi_\alpha^k d\mu$  are well-defined for  $b \in \text{BMO}(\mu)$ . The following injectivity property is somewhat technical, and we postpone its proof to Appendix A:

**11.3. Proposition.** *Suppose that  $b \in \text{BMO}(\mu)$  satisfies  $(b, \psi_\alpha^k) = 0$  for all  $k \in \mathbb{Z}$  and  $\alpha \in \mathcal{Y}^k$ . Then  $b$  is equal to a constant.*

If  $X$  is bounded, one should mention that  $k \geq k_0$ , but in fact, notice that  $\mathcal{Y}^k = \emptyset$  and so  $\psi_\alpha^k$  does not even exist when  $k < k_0$ . We thus do not need to distinguish further between  $X$  bounded or not.

Note that some nontrivial a priori size condition on  $b$  is in general necessary for such a conclusion. For example, if the  $\psi_\alpha^k$  are regular wavelets on  $\mathbb{R}^d$ , then  $(\psi_\alpha^k, P) = 0$  for all polynomials  $P$  of degree lower than the regularity of the wavelets.

We say that a sequence  $\{b_\alpha^k\}_{k \in \mathbb{Z}, \alpha \in \mathcal{Y}^k}$  is a **Carleson sequence** if

$$\|\{b_\alpha^k\}_{k, \alpha}\|_{\text{Car}} := \sup_{\ell \in \mathbb{Z}, \beta \in \mathcal{X}^\ell} \left( \frac{1}{\mu(Q_\beta^\ell)} \sum_{\substack{k \in \mathbb{Z}, \alpha \in \mathcal{Y}^k \\ (k+1, \alpha) \leq (\ell, \beta)}} |b_\alpha^k|^2 \right)^{1/2} < \infty.$$

Pay attention to the fact that the supremum runs over all  $\ell$  and  $\beta \in \mathcal{X}^\ell$ , which index the dyadic cubes  $Q_\beta^\ell$ , whereas the sum runs over  $k$  and  $\alpha \in \mathcal{Y}^k = \mathcal{X}^{k+1} \setminus \mathcal{X}^k$ . Via the wavelet decomposition, we obtain an isomorphism between BMO functions and Carleson sequences:

**11.4. Theorem.** *The spaces  $\text{BMO}(\mu)/\mathbb{C}$  (BMO functions modulo constants) and  $\text{Car}$  are isomorphic. This isomorphism is realized via  $b \mapsto \{(b, \psi_\alpha^k)\}_{k, \alpha}$ , with inverse given by*

$$(11.5) \quad \{b_\alpha^k\}_{k, \alpha} \mapsto \sum_{k, \alpha} b_\alpha^k \left( \psi_\alpha^k - 1_{\{j: \delta^j > r_0\}}(k) \psi_\alpha^k(x_0) \right),$$

where the series converges in  $L_{\text{loc}}^2(\mu)$  for every  $x_0 \in X$  and  $r_0 > 0$ , and the choices of  $x_0$  and  $r_0$  only alter the result by an additive constant.

The result and its proof are reasonably classical in spirit, but the lack of bounded support of the wavelets somewhat complicates the matters. While exponential decay is intuitively almost as good, one needs to go through certain technicalities if one wants to be careful with convergence issues. We indicate the argument.

Note that we have formulated the recovery of the BMO function from its wavelet coefficients using the “infra-red” renormalisation of the wavelet series (since the modification on the series appears for large scales which corresponds to small frequencies in the classical Euclidean situation) rather than the  $H^1$ –BMO duality; this is in contrast to the statement in Meyer’s book [30], for instance, but this other method would be certainly doable here as well.

*Proof.* Let us first see that  $b \mapsto \{(b, \psi_\alpha^k)\}_{k, \alpha}$  maps  $\text{BMO}(\mu)/\mathbb{C}$  into Carleson sequences. The injectivity of this mapping is the content of Proposition 11.3. Given  $(k, \alpha)$ , we let  $\tilde{B} = B(x_\alpha^k, C\delta^k)$ , with some large  $C$ , be a ball such that  $d(Q_\alpha^k, (B_\alpha^k)^c) \gtrsim \delta^k$ . We write

$$b = (b - b_{\tilde{B}})1_{\tilde{B}} + (b - b_{\tilde{B}})1_{\tilde{B}^c} + b_{\tilde{B}} =: b^1 + b^2 + b^3.$$

Then

$$\sum |(b^1, \psi_\beta^\ell)|^2 \leq \|b^1\|_2^2 \lesssim \|b\|_{\text{BMO}(\mu)} \mu(Q_\alpha^k)$$

by orthogonality, John–Nirenberg inequality and doubling;  $(b^2, \psi_\beta^\ell)$  may be estimated by the decay of the wavelets, and  $(b^3, \psi_\alpha^k) = 0$ .

*From Carleson sequences to BMO.* Conversely, assume the Carleson condition. Given a ball  $B_1 = B(x_1, r_1)$ , we may rearrange  $1_{B_1}$  times the right side of (11.5) as

$$\begin{aligned} & 1_{B_1} \sum_{k:\delta^k \leq r_1} \sum_{\alpha: y_\alpha^k \in CB_1} b_\alpha^k \psi_\alpha^k + 1_{B_1} \sum_{k:\delta^k \leq r_1} \sum_{\alpha: y_\alpha^k \in (CB_1)^c} b_\alpha^k \psi_\alpha^k + 1_{B_1} \sum_{k:\delta^k > r_1} \sum_{\alpha} b_\alpha^k (\psi_\alpha^k - \psi_\alpha^k(x_1)) \\ & - 1_{B_1} \sum_{k:r_0 < \delta^k \leq r_1} \sum_{\alpha} b_\alpha^k \psi_\alpha^k(x_0) + 1_{B_1} \sum_{k:r_1 < \delta^k \leq r_0} \sum_{\alpha} b_\alpha^k \psi_\alpha^k(x_1) \\ & + 1_{B_1} \sum_{k:\delta^k > r_0 \vee r_1} b_\alpha^k (\psi_\alpha^k(x_1) - \psi_\alpha^k(x_0)). \end{aligned}$$

Here the last three terms converge uniformly to  $1_{B_1}$  times a constant, the first term converges in  $L^2(\mu)$  with norm bounded by  $\mu(B_1)^{1/2}$  (use the Carleson condition after covering  $CB_1$  by boundedly many dyadic cubes of sidelength  $\sim r_1$ ), and the second and third terms can be estimated uniformly by the decay and regularity of the wavelets  $\psi_\alpha^k$ . This proves the  $L^2(\mu)$  convergence on  $B_1$ , and also the BMO estimate

$$\|1_{B_1}(\tilde{b} - c_{B_1})\|_2 \lesssim \mu(B_1)^{1/2},$$

where  $\tilde{b}$  stands for the function on the right of (11.5), and  $c_{B_1}$  is the constant produced by the last three terms in the above expansion. Effectively, the same argument also shows the possibility of replacing  $(x_0, r_0)$  in (11.5) by  $(x_1, r_1)$ , only changing the result by a constant.

*Verifying that the two mappings are inverses to each other.* This is the most technical part of the argument. For a Carleson sequence  $\{b_\alpha^k\}_{k,\alpha}$ , let  $\tilde{b}$  denote the  $\text{BMO}(\mu)$  function (as shown in the previous part) on the right of (11.5). We claim that  $(\tilde{b}, \psi_\alpha^k) = b_\alpha^k$ , which completes the proof. Indeed, this also shows that if  $\{b_\alpha^k\}_{k,\alpha}$  arose as  $b_\alpha^k = (b, \psi_\alpha^k)$  from some  $b \in \text{BMO}(\mu)$ , then the function  $\tilde{b}$  produced by (11.5) satisfied  $(\tilde{b} - b, \psi_\alpha^k) = 0$  and hence  $\tilde{b} = b + \text{constant}$ .

While the claim is formally obvious, due to the orthogonality

$$(\psi_\beta^\ell - 1_{\{j:\delta^j > r_0\}}(\ell) \psi_\beta^\ell(x_0), \psi_\alpha^k) = \delta_{k\ell} \delta_{\alpha\beta},$$

we need to justify exchanging the order of summation and integration. To this end, let us reinvestigate the convergence of (11.5) where, since  $\psi_\alpha^k$  is orthogonal to constants, we may assume that  $(x_0, r_0) = (y_\alpha^k, \delta^k)$ . Then

$$\tilde{b} = \sum_{\theta} \sum_{(\ell+1, \beta) \leq (k+1, \theta)} b_\beta^\ell \psi_\beta^\ell + \sum_{\ell:\delta^\ell > \delta^k} \sum_{\beta} b_\beta^\ell (\psi_\beta^\ell - \psi_\beta^\ell(y_\alpha^k)) =: \tilde{b}^1 + \tilde{b}^2.$$

The second part converges pointwise absolutely to a limit of size  $\lesssim 1 + \log_+(d(x, y_\alpha^k)/\delta^k)$ . Since  $\psi_\alpha^k$  decays exponentially, this allows to exchange the summation and integration when computing  $(\tilde{b}^2, \psi_\alpha^k)$ . Since  $\psi_\alpha^k$  is orthogonal to  $\psi_\beta^\ell$  for  $\delta^\ell > \delta^k$ , as well as to constants, we get  $(\tilde{b}^2, \psi_\alpha^k) = 0$ .

For the first part, we have

$$\tilde{b}^1 =: \sum_{\theta} \tilde{b}_\theta = \sum_{\theta} \sum_{\zeta} 1_{Q_\zeta^{k+1}} \tilde{b}_\theta, \quad \|1_{Q_\zeta^{k+1}} \tilde{b}_\theta\|_2 \lesssim \sqrt{\mu(Q_\zeta^{k+1})} \exp(-\gamma(\delta^{-k} d(x_\zeta^{k+1}, x_\theta^{k+1}))^a),$$

whereas

$$\|1_{Q_\zeta^{k+1}} \psi_\alpha^k\|_2 \lesssim \exp(-\gamma(\delta^{-k} d(x_\zeta^{k+1}, x_\alpha^{k+1}))^a).$$



Hence

$$\begin{aligned}
\sum_{\theta} \|\tilde{b}_{\theta} \psi_{\alpha}^k\|_1 &= \sum_{\theta} \sum_{\zeta} \|1_{Q_{\zeta}^{k+1}} \tilde{b}_{\theta} \psi_{\alpha}^k\|_1 \\
&\lesssim \sum_{\zeta} \sum_{\theta} \sqrt{\mu(Q_{\zeta}^{k+1})} \exp(-\gamma(\delta^{-k} d(x_{\zeta}^{k+1}, x_{\theta}^{k+1}))^a) \exp(-\gamma(\delta^{-k} d(x_{\zeta}^{k+1}, x_{\alpha}^{k+1}))^a) \\
&\lesssim \sum_{\zeta} \sqrt{\mu(Q_{\zeta}^{k+1})} \exp(-\gamma(\delta^{-k} d(x_{\zeta}^{k+1}, x_{\alpha}^{k+1}))^a) \lesssim \sqrt{\mu_{\alpha}^{k+1}}.
\end{aligned}$$

This allows the first exchange in

$$(\tilde{b}^1, \psi_{\alpha}^k) = \sum_{\theta} (\tilde{b}_{\theta}, \psi_{\alpha}^k) = \sum_{\theta} \sum_{(\ell+1, \beta) \leq (k+1, \theta)} b_{\beta}^{\ell}(\psi_{\beta}^{\ell}, \psi_{\alpha}^k) = \sum_{\theta} \sum_{(\ell+1, \beta) \leq (k+1, \theta)} b_{\beta}^{\ell} \delta_{\ell k} \delta_{\alpha \beta} = b_{\alpha}^k,$$

where the second exchange follows from the convergence of  $\sum_{(\ell+1, \beta) \leq (k+1, \theta)} b_{\beta}^{\ell} \psi_{\beta}^{\ell}$  in  $L^2(\mu)$ , and from  $\psi_{\alpha}^k \in L^2(\mu)$ . Altogether, we get  $(\tilde{b}, \psi_{\alpha}^k) = b_{\alpha}^k$ , as claimed.  $\square$

## 12. THE $T(1)$ THEOREM

To illustrate the power of the spline wavelets, we use them to sketch a proof of the  $T(1)$  theorem in any space of homogeneous type. Such a result is surely part of the folklore, but surprisingly difficult to find spelled out in complete generality: the seminal paper of David–Journé–Semmes [9] makes several assumptions on the space, like the small boundary property of balls, and many recent references treat other special cases like Ahlfors–David [11] or reverse doubling spaces [15].

The technology to prove the  $T(1)$  theorem in a general space of homogeneous type has certainly existed since the work of M. Christ [5]. Indeed, a proof of the  $T(1)$  theorem can be given by using the Haar wavelets only, and these have been available since Christ’s construction of his dyadic cubes with the small boundary property. In fact, Christ even formulates the general  $T(1)$  theorem [5, Theorem 8], but attributes it to [9], and proceeds to use it as a tool for proving a certain ‘local’ variant. However, Christ’s techniques would have clearly delivered a proof of the ‘global’  $T(1)$  theorem as well, without the restrictions imposed in [9].

So the regular wavelets are not strictly necessary for obtaining the  $T(1)$  theorem, but they nevertheless provide a rather efficient tool for that purpose. We only indicate the argument, which largely imitates the treatment in Euclidean cases given by Meyer [30].

We take as space of test functions the space  $\mathcal{V}_s = C_0^s(X)$  of functions with bounded support and Hölder-regularity  $s$  equipped with the usual topology, where  $s \in (0, \eta)$  is arbitrary and we recall that  $\eta$  is the regularity of the splines. The space  $\mathcal{V}_s$  is dense in  $L^2(\mu)$  by Proposition 4.5. Let  $\mathcal{V}'_s$  denotes its dual space. Recall the standard definition.

**12.1. Definition.** Let  $T : \mathcal{V}_s \rightarrow \mathcal{V}'_s$  be a linear continuous operator. We say that  $T$  is associated to a Calderón-Zygmund kernel of order  $s$  if the distributional kernel  $K(x, y)$  of  $T$  satisfies for some constant  $C_1 < \infty$ ,

$$|K(x, y)| \lesssim C_1 V(x, y)^{-1}$$

when  $x \neq y$  and

$$|K(x, y) - K(x, y')| \leq C_1 \left( \frac{d(y, y')}{d(x, y)} \right)^s V(x, y)^{-1}$$

when  $x, y, y' \in X$  with  $0 < d(x, x') \leq d(x, y)/(2A_0)$

$$|K(x, y) - K(x, y')| \leq C_1 \left( \frac{d(x, x')}{d(x, y)} \right)^s V(x, y)^{-1}$$

when  $x, x', y \in X$  with  $0 < d(x, x') \leq d(x, y)/(2A_0)$  and if furthermore, for any  $f \in \mathcal{V}_s$ , one has the representation

$$(12.2) \quad Tf(x) = \int K(x, y) f(y) d\mu(y)$$

for almost every  $x \notin \text{supp } f$ .

Recall also the weak boundedness property and the meaning of  $T(1)$ . A linear continuous operator  $T : \mathcal{V}_s \rightarrow \mathcal{V}'_s$  has weak boundedness property  $WBP(\sigma)$  if

$$|(Tf, g)| \leq C_0 V(x, r)$$

whenever  $f, g \in \mathcal{V}_\sigma$ , with support in  $B(x, r)$  and are normalized by  $\|f\|_\infty + r^\sigma \|f\|_{\dot{C}^\sigma} \leq 1$  and similarly for  $g$ . It is well-known that  $WBP(\sigma)$  and  $WBP(\sigma')$  are equivalent conditions whenever  $s \leq \sigma, \sigma' \leq \eta$  when  $T$  is associated to a Calderón-Zygmund kernel of order  $s$ . As for  $T(1)$ , it is defined as a continuous linear functional on the subspace of  $\mathcal{V}_s$  of functions  $f$  with mean value 0 by

$$(T(1), f) = (Tg, f) + \int_X (1 - g(x)) ({}^t T f)(x) d\mu(x)$$

with  $g$  a function in  $\mathcal{V}_s$  that is 1 on a ball  $B(x_0, r)$  containing the support of  $f$  and 0 on the complement of  $B(x_0, 2A_0 r)$  as in Corollary 4.2, as  ${}^t T f$  is integrable away the support of  $f$ .

**12.3. Theorem.** *Let  $(X, d, \mu)$  be any space of homogeneous type and  $T$  be associated to a Calderón-Zygmund kernel of order  $s$ . Then  $T$  has a bounded extension to  $L^2(\mu)$  if and only if  $T$  has  $WBP(s)$ ,  $T(1) \in BMO(\mu)$ ,  ${}^t T(1) \in BMO(\mu)$ .*

As usual it suffices to prove the converse. With our spline wavelets at hand, any of the standard wavelet proofs of the  $T(1)$  theorem applies to our statement. For example, one can follow almost line by line the wavelet proof given in [30, pp. 267-270]. The weak boundedness property and the kernel representation allows to define  $(T(1), \psi_\alpha^k)$ , whose absolute value has a bound  $C\sqrt{\mu(B(y_\alpha^k, \delta^k))}$ . Then, one takes away the paraproducts and set  $U = T - \Pi_{T(1)} - {}^t \Pi_{T(1)}$  where

$$\Pi_\beta f = \sum_{(k, \alpha)} (f, \bar{s}_\alpha^{k+1})(\beta, \psi_\alpha^k) \psi_\alpha^k$$

with  $\bar{s}_\alpha^{k+1}$  the spline  $s_\alpha^{k+1}$  normalized in  $L^1(\mu)$ . Recall that  $\psi_\alpha^k$  is localized near the point  $y_\alpha^k = x_\alpha^{k+1} \in \mathcal{Y}^k = \mathcal{X}^{k+1} \setminus \mathcal{X}^k$ , so that the Calderón-Zygmund estimates follow from the results in Section 9 (see the initial comment there) when  $|(\beta, \psi_\alpha^k)| \leq C\sqrt{\mu(B(y_\alpha^k, \delta^k))}$ . The paraproduct is classically bounded on  $L^2(\mu)$  if and only if the coefficients  $(\beta, \psi_\alpha^k)$  form a Carleson sequence. It uses the maximal function  $f^*(x) = \sup |(f, \bar{s}_\alpha^{k+1})|$  taken over all  $k \in \mathbb{Z}, x_\alpha^{k+1} \in \mathcal{X}^{k+1}$  such that  $x \in B(x_\alpha^{k+1}, C\delta^{k+1})$  which is clearly comparable to Hardy-Littlewood maximal function. By Theorem 11.4, this is equivalent to  $\beta \in BMO(\mu)/\mathbb{C}$ . Thus  $U$  is associated to a Calderón-Zygmund kernel of order  $s$  with, moreover,  $U(1) = {}^t U(1) = 0$ .

Next, by taking  $f, g$  as finite linear combination of the wavelets (which are  $L^2$  dense), it suffices to prove that the coefficients of  $U$  on the wavelet basis form a bounded matrix on  $\ell^2$ . Here, using only kernel regularity estimates (not the size, in fact), weak boundedness,  $U(1) = {}^t U(1) = 0$ , one makes sense of the coefficients  $(U\psi_\alpha^k, \psi_\beta^\ell)$  and finds the estimate

$$(12.4) \quad |(U\psi_\alpha^k, \psi_\beta^\ell)| \lesssim \frac{C_0 \delta^{|k-\ell|\varepsilon} (1 + \delta^{-k \wedge \ell} d(y_\alpha^k, y_\beta^\ell))^{-\varepsilon} \sqrt{\mu(B(y_\alpha^k, \delta^k)) \mu(B(y_\beta^\ell, \delta^\ell))}}{\mu(B(y_\alpha^k, \delta^k)) + \mu(B(y_\beta^\ell, \delta^\ell)) + V(y_\alpha^k, y_\beta^\ell)}$$

with  $0 < \varepsilon < s$  and  $k \wedge \ell = \inf(k, \ell)$ . The version of the Schur lemma for the  $\ell^2$  boundedness is

$$\sum_{(k, \alpha)} |(U\psi_\alpha^k, \psi_\beta^\ell)| \sqrt{\mu(B(y_\alpha^k, \delta^k))} \lesssim \sqrt{\mu(B(y_\beta^\ell, \delta^\ell))}$$

uniformly in  $(\ell, \beta)$  and the symmetric estimate reversing the roles of  $(k, \alpha), (\ell, \beta)$ . Details are left to the reader.

Strictly speaking, this argument works when  $\mu(X) = \infty$ . When  $\mu(X) < \infty$ , one has to incorporate the following observation. First, the assumption that  $T(1) \in BMO(\mu)$  implies in particular that  $T(1)$  is (identified to) a locally integrable function. Second, since  $X$  is bounded, it can be regarded as a ball, and the constant function 1 is a smooth bump function associated with this ball. Hence, the weak boundedness property implies that  $|\int_X T(1) d\mu| = |\langle T(1), 1 \rangle| \lesssim \mu(X)$ , and thus  $\|T(1)\|_{BMO(\mu)} + |\int_X T(1) d\mu| < \infty$  (recall that our BMO norm is the homogeneous norm). By John-Nirenberg's inequality, this implies that  $T(1) \in L^2(\mu)$  with  $\|T(1)\|_2 \lesssim \|T(1)\|_{BMO(\mu)} +$

$|\int_X T(1) d\mu|$ . Similarly  $\|{}^tT(1)\|_2 \lesssim \|{}^tT(1)\|_{\text{BMO}(\mu)} + |\int_X {}^tT(1) d\mu| < \infty$ . The argument before can be repeated and implies that  $\pi T \pi$  is bounded where  $\pi$  is the orthogonal projection onto the subspace of functions in  $L^2(\mu)$  with mean value 0 since the wavelets span this space. The boundedness of  $T$  on  $L^2(\mu)$  follows readily.

**12.5. Remark.** As in [30], estimate (12.4) is stable under matrix multiplication up to changing  $\varepsilon$  to a smaller value. This algebra property furnishes a proof that Calderón–Zygmund operators with  $T(1) = {}^tT(1) = 0$  is an algebra for the composition. This was proved in [13] by working with a new Hölder-continuous quasi-distance as in [9], hence changing the class of singular integrals as discussed in the Introduction.

### 13. REDUNDANCY OF THE SIZE ESTIMATE

It turns out that the size condition on the kernel is actually redundant in the  $T(1)$  theorem, in that it already follows from regularity of the kernel and the weak boundedness property. To our knowledge, this remark seems new even in the context of  $\mathbb{R}^d$  with the Lebesgue measure. We came across this observation noticing that the regularity estimate of kernels like  $\sum h_\alpha^k(x) h_\alpha^k(y)$  with  $h_\alpha^k$  satisfying the size and regularity of spline or spline-wavelets is straightforward, while the size estimate required the analysis in Section 9 based on the structure of the point sets  $\mathcal{Y}^k$ . It is also possible to develop the  $T(1)$  theory without using the size estimate at all.

To be more precise, let us say that  $T : \mathcal{V}_s \rightarrow \mathcal{V}'_s$  is associated to a Calderón–Zygmund kernel of order  $s$  in the relaxed sense, if in Definition 12.1 we leave out the condition that  $|K(x, y)| \leq C_1 V(x, y)^{-1}$ , and only assume that  $K(x, y)$  is locally integrable away from the diagonal. This condition could be further relaxed, by not assuming the a priori existence of a measurable kernel at all, only that  $T$  is a weak limit of operators  $T_n$ , which are associated to relaxed kernels of order  $s$  in a uniform way. However, we stick to the stated relaxation, for the simplicity of formulation.

**13.1. Proposition.** *Suppose that  $T : \mathcal{V}_s \rightarrow \mathcal{V}'_s$  is associated to a Calderón–Zygmund kernel of order  $s$  in the relaxed sense, and that  $T$  satisfies WBP( $\sigma$ ) for some  $\sigma \in [s, \eta]$ . Then  $T$  is actually associated to a Calderón–Zygmund kernel of order  $s$  in the sense of Definition 12.1.*

*Proof.* Recall that the paraproducts  $\Pi_{T(1)}$  and  ${}^t\Pi_{T(1)}$  are associated to Calderón–Zygmund kernels of order  $s$  under these assumptions. Thus it suffices to prove the claim for the operator  $U = T - \Pi_{T(1)} - {}^t\Pi_{T(1)}$  in place of  $T$ . The kernel of  $U$  is given by

$$K(x, y) = \sum_{k, \ell, \alpha, \beta} (U \psi_\alpha^k, \psi_\beta^\ell) \psi_\alpha^k(y) \psi_\beta^\ell(x),$$

where the coefficients satisfy the estimate (12.4), as this estimate did not make use of the size of the kernel.

We show that this series converges absolutely and satisfies the size estimate. By symmetry, it suffices to consider the half of the sum with  $\ell \leq k$ , hence  $\delta^\ell \geq \delta^k$ , and

$$\begin{aligned} & |(U \psi_\alpha^k, \psi_\beta^\ell) \psi_\alpha^k(y) \psi_\beta^\ell(x)| \\ & \lesssim \frac{\delta^{(k-\ell)\varepsilon} (1 + \delta^{-\ell} d(y_\alpha^k, x_\beta^\ell))^{-\varepsilon}}{V(y_\beta^\ell, \delta^\ell) + V(y_\alpha^k, y_\beta^\ell)} \exp(-\gamma(\delta^{-k} d(y, y_\alpha^k))^a) \exp(-\gamma(\delta^{-k} d(x, y_\beta^\ell))^a) \\ & \lesssim \frac{\delta^{(k-\ell)\varepsilon} (1 + \delta^{-\ell} d(y, x))^{-\varepsilon}}{V(x, \delta^\ell) + V(y, x)} \exp(-\gamma(\delta^{-k} d(y, y_\alpha^k))^a) \exp(-\gamma(\delta^{-\ell} d(x, y_\beta^\ell))^a) \end{aligned}$$

Hence

$$\begin{aligned} & \sum_{k, \ell: \ell \leq k} \sum_{\alpha, \beta} |(U \psi_\alpha^k, \psi_\beta^\ell) \psi_\alpha^k(y) \psi_\beta^\ell(x)| \\ & \lesssim \sum_{k, \ell: \ell \leq k} \frac{\delta^{(k-\ell)\varepsilon} (1 + \delta^{-\ell} d(y, x))^{-\varepsilon}}{V(x, \delta^\ell) + V(y, x)} \exp(-\gamma(\delta^{-k} d(y, \mathcal{Y}^k))^a) \exp(-\gamma(\delta^{-\ell} d(x, \mathcal{Y}^\ell))^a) \\ & \lesssim \sum_{\ell: \delta^\ell \geq d(x, y)} \sum_{k: \delta^k \leq \delta^\ell} \frac{\delta^{(k-\ell)s}}{V(x, \delta^\ell)} \exp(-\gamma(\delta^{-\ell} d(x, \mathcal{Y}^\ell))^a) + \sum_{k, \ell: d(x, y) \geq \delta^\ell \geq \delta^k} \frac{\delta^{ks} d(x, y)^{-s}}{V(x, y)}, \end{aligned}$$

where  $(1 + \delta^{-\ell} d(x, y))^{-s}$  was dominated by 1 in the first sum, and by  $(\delta^{-\ell} d(x, y))^{-s} = \delta^{\ell s} d(x, y)^{-s}$  in the second. In the first part, we sum a geometric series in  $k$ , and then use the bound of Lemma 8.3. In the second part, we simply sum up a geometric series in  $k$  producing the factor  $\delta^{\ell s}$ , and then a geometric series in  $\ell$ .  $\square$

### APPENDIX A. PROOF OF PROPOSITION 11.3

We begin with a technical estimate:

**A.1. Lemma.** *For a any ball  $B = B(x_0, r)$  and  $b \in \text{BMO}(\mu)$  of norm one, we have*

$$|(\psi_\alpha^k, 1_{B^c}(b - b_B))| \lesssim \sqrt{\mu(B_\alpha^k)} \times \begin{cases} 1 + \log((\delta^k + r + d(y_\alpha^k, x_0))/\min\{\delta^k, r\}) & \text{always,} \\ e^{-\gamma(r/\delta^k)^a} & \text{if } y_\alpha^k \in \frac{1}{2A_0}B \text{ and } \delta^k \leq r, \end{cases}$$

with  $B_\alpha^k := B(y_\alpha^k, \delta^k)$ .

*Proof.* By estimating  $|\psi_\alpha^k| \leq C\mu(B_\alpha^k)^{-1/2}e^{-\gamma\delta^{-ma}}$  on  $\delta^{-m}B_\alpha^k \setminus \delta^{-(m-1)}B_\alpha^k$ , and simply ignoring the indicator, we estimate up by

$$(|\psi_\alpha^k|, |b - b_B|) \leq \frac{C}{\sqrt{\mu(B_\alpha^k)}} \sum_{m=0}^{\infty} e^{-\gamma\delta^{-ma}} \int_{\delta^{-m}B_\alpha^k} |b - b_B| d\mu,$$

where

$$\begin{aligned} \int_{\delta^{-m}B_\alpha^k} |b - b_B| d\mu &\leq \int_{\delta^{-m}B_\alpha^k} |b - b_{\delta^{-m}B_\alpha^k}| d\mu + \mu(\delta^{-m}B_\alpha^k) |b_{\delta^{-m}B_\alpha^k} - b_B| \\ &\lesssim \mu(\delta^{-m}B_\alpha^k) \left(1 + \log \frac{\delta^{-m}\delta^k + r + d(y_\alpha^k, x_0)}{\min\{\delta^{-m}\delta^k, r\}}\right) \\ &\lesssim \delta^{-mM} \mu(B_\alpha^k) \left(1 + m + \log \frac{\delta^k + r + d(y_\alpha^k, x_0)}{\min\{\delta^k, r\}}\right), \end{aligned}$$

and hence

$$(|\psi_\alpha^k|, |b - b_B|) \leq C\sqrt{\mu(B_\alpha^k)} \left(1 + \log \frac{\delta^k + r + d(y_\alpha^k, x_0)}{\min\{\delta^k, r\}}\right).$$

If  $y_\alpha^k \in \frac{1}{2A_0}B$ , so that  $d(y_\alpha^k, B^c) \gtrsim r$ , and  $\delta^k \leq r$ , we make a similar estimate but exploiting the fact that for  $\delta^{-m}B_\alpha^k$  to touch  $B^c$ , we need  $\delta^{-m}\delta^k \gtrsim r(B)$ . Thus

$$\begin{aligned} (|\psi_\alpha^k|, 1_{B^c}|b - b_B|) &\leq \frac{C}{\sqrt{\mu(B_\alpha^k)}} \sum_{m: \delta^{-m} \gtrsim r/\delta^k} e^{-\gamma\delta^{-ma}} \int_{\delta^{-m}B_\alpha^k} |b - b_B| d\mu, \\ &\leq \frac{C}{\sqrt{\mu(B_\alpha^k)}} \sum_{m: \delta^{-m} \gtrsim r/\delta^k} e^{-\gamma\delta^{-ma}} \mu(\delta^{-m}B_\alpha^k) \left(1 + m + \log \frac{r}{\delta^k}\right) \\ &\leq C\sqrt{\mu(B_\alpha^k)} \cdot e^{-\gamma(r/\delta^k)^a}. \end{aligned} \quad \square$$

We are ready for the proof of Proposition 11.3, which we recall here:

**A.2. Proposition.** *Suppose that  $b \in \text{BMO}(\mu)$  satisfies  $(\psi_\alpha^k, b) = 0$  for all  $k, \alpha$ . Then  $b$  is equal to a constant.*

*Proof.* We show for every ball  $B = B(x_0, r)$  and  $\varepsilon > 0$  the following: there exists a constant  $c$  such that  $\|1_B(b - c)\|_\infty \leq \varepsilon$ . This clearly suffices. We may assume for simplicity that  $\|b\|_{\text{BMO}(\mu)} \leq 1$ .

Given  $B$ , we take a large auxiliary  $\tilde{B} = B(x_0, \tilde{r})$ . We use the fact that  $1_{\tilde{B}}(b - b_{\tilde{B}}) \in L^2(\mu)$  can be expanded in terms of the wavelets  $\psi_\alpha^k$ :

$$\begin{aligned}
 1_B b &= 1_B 1_{\tilde{B}}(b - b_{\tilde{B}}) + 1_B b_{\tilde{B}} = 1_B \sum_{k, \alpha} \psi_\alpha^k(\psi_\alpha^k, 1_{\tilde{B}}(b - b_{\tilde{B}})) + 1_B b_{\tilde{B}} \\
 (A.3) \quad &= 1_B \sum_{k: \delta^k \leq r'} \sum_{\alpha} \psi_\alpha^k(\psi_\alpha^k, 1_{\tilde{B}}(b - b_{\tilde{B}})) + 1_B \sum_{k: \delta^k > r'} \sum_{\alpha} (\psi_\alpha^k - \psi_\alpha^k(x_0))(\psi_\alpha^k, 1_{\tilde{B}}(b - b_{\tilde{B}})) \\
 &\quad + 1_B \sum_{k: \delta^k > r'} \sum_{\alpha} \psi_\alpha^k(x_0)(\psi_\alpha^k, 1_{\tilde{B}}(b - b_{\tilde{B}})) + 1_B b_{\tilde{B}}.
 \end{aligned}$$

Provided that everything converges (which we check in a moment), the terms on the last line give  $1_B$  times a constant, so it suffices to show that the second-to-last line becomes arbitrarily small for properly chosen  $\tilde{r} > r' > r$ . For the convergence of the last line, note that

$$\begin{aligned}
 |\psi_\alpha^k(x_0)(\psi_\alpha^k, 1_{\tilde{B}}(b - b_{\tilde{B}}))| &\leq |\psi_\alpha^k(x_0)| \|\psi_\alpha^k\|_\infty \|1_{\tilde{B}}(b - b_{\tilde{B}})\|_1 \\
 &\leq \frac{C}{\mu(B_\alpha^k)} \exp(-\gamma(\delta^{-k}d(x_0, y_\alpha^k))^a) \mu(\tilde{B}),
 \end{aligned}$$

and, by Lemma 8.3,

$$\begin{aligned}
 &\sum_{k: \delta^k > r'} \sum_{\alpha} \frac{C}{\mu(B_\alpha^k)} \exp(-\gamma(\delta^{-k}d(x_0, y_\alpha^k))^a) \\
 &\leq \sum_{k: \delta^k > r'} \frac{C}{V(x_0, \delta^k)} \exp(-\gamma(\delta^{-k}d(x_0, \mathcal{Y}^k))^a) \leq \frac{C}{V(x_0, r')}.
 \end{aligned}$$

So the last line of (A.3) is a well-defined constant, as claimed, and it remains to see that the rest of the right side of the same equation is small.

Part  $\delta^k \leq r'$ . Since  $\psi_\alpha^k$  is orthogonal to constants always, and to  $b$  by assumption, we have

$$(\psi_\alpha^k, 1_{\tilde{B}}(b - b_{\tilde{B}})) = -(\psi_\alpha^k, 1_{\tilde{B}^c}(b - b_{\tilde{B}})),$$

and we may apply Lemma A.1 (with  $\tilde{B}$  in place of  $B$ ) to estimate. If  $y_\alpha^k \in (\frac{1}{2A_0}\tilde{B})^c$ , then

$$|(\psi_\alpha^k, 1_{\tilde{B}^c}(b - b_{\tilde{B}}))| \lesssim \sqrt{\mu(B_\alpha^k)} \left(1 + \log \frac{d(y_\alpha^k, x_0)}{\delta^k}\right),$$

whereas

$$\|1_B \psi_\alpha^k\|_\infty \lesssim \frac{1}{\sqrt{\mu(B_\alpha^k)}} \exp(-\gamma(\delta^{-k}d(x_0, y_\alpha^k))^a).$$

Hence

$$\begin{aligned}
 (A.4) \quad &\sum_{k: \delta^k \leq r'} \sum_{\alpha: y_\alpha^k \in (\frac{1}{2A_0}\tilde{B})^c} \|1_B \psi_\alpha^k(\psi_\alpha^k, 1_{\tilde{B}}(b - b_{\tilde{B}}))\|_\infty \lesssim \sum_{k: \delta^k \leq r'} \sum_{\alpha: y_\alpha^k \in (\frac{1}{2A_0}\tilde{B})^c} \exp(-\gamma(\delta^{-k}d(x_0, y_\alpha^k))^a) \\
 &\lesssim \sum_{k: \delta^k \leq r'} e^{-\gamma(\tilde{r}/\delta^k)^a} \lesssim e^{-\gamma(\tilde{r}/r')^a}.
 \end{aligned}$$

If, on the other hand,  $y_\alpha^k \in \frac{1}{2A_0}\tilde{B}$ , then we have the estimates

$$|(\psi_\alpha^k, 1_{\tilde{B}^c}(b - b_{\tilde{B}}))| \lesssim \sqrt{\mu(B_\alpha^k)} \cdot e^{-\gamma(\tilde{r}/\delta^k)^a}, \quad \|1_B \psi_\alpha^k\|_\infty \lesssim \frac{1}{\sqrt{\mu(B_\alpha^k)}},$$

hence

$$\begin{aligned}
 (A.5) \quad &\sum_{k: \delta^k \leq r'} \sum_{\alpha: y_\alpha^k \in \frac{1}{2A_0}\tilde{B}} \|1_B \psi_\alpha^k(\psi_\alpha^k, 1_{\tilde{B}}(b - b_{\tilde{B}}))\|_\infty \lesssim \sum_{k: \delta^k \leq r'} \sum_{\alpha: y_\alpha^k \in \frac{1}{2A_0}\tilde{B}} e^{-\gamma(\tilde{r}/\delta^k)^a} \\
 &\lesssim \sum_{k: \delta^k \leq r'} e^{-\gamma(\tilde{r}/\delta^k)^a} (\tilde{r}/\delta^k)^M \lesssim e^{-\gamma(\tilde{r}/r')^a}.
 \end{aligned}$$

Part  $\delta^k > r'$ . As before, we have from Lemma A.1 (with  $\tilde{B}$  in place of  $B$ ) that

$$\begin{aligned} |(\psi_\alpha^k, 1_{\tilde{B}}(b - b_{\tilde{B}}))| &\lesssim \sqrt{\mu(B_\alpha^k)} \left(1 + \log \frac{\delta^k + \tilde{r} + d(y_\alpha^k, x_0)}{\min\{\delta^k, \tilde{r}\}}\right) \\ &\lesssim \sqrt{\mu(B_\alpha^k)} \left(1 + \log_+ \frac{\delta^k}{\tilde{r}} + \log_+ \frac{\tilde{r}}{\delta^k} + \log_+ \frac{d(y_\alpha^k, x_0)}{\delta^k}\right), \end{aligned}$$

and from the regularity of the wavelets that

$$\|1_B(\psi_\alpha^k - \psi_\alpha^k(x_0))\|_\infty \lesssim \frac{1}{\sqrt{\mu(B_\alpha^k)}} \left(\frac{r}{\delta^k}\right)^\eta \exp(-\gamma(\delta^{-k}d(y_\alpha^k, x_0))^a).$$

Hence

$$\begin{aligned} (A.6) \quad &\sum_{k:\delta^k > r'} \sum_{\alpha} \|1_B(\psi_\alpha^k - \psi_\alpha^k(x_0))(\psi_\alpha^k, 1_{\tilde{B}}(b - b_{\tilde{B}}))\|_\infty \\ &\lesssim \sum_{k:\delta^k > r'} \left(\frac{r}{\delta^k}\right)^\eta \left(1 + \log_+ \frac{\tilde{r}}{\delta^k} + \log_+ \frac{\delta^k}{\tilde{r}}\right) \\ &\quad \times \sum_{\alpha} \left(1 + \log_+ \frac{d(y_\alpha^k, x_0)}{\delta^k}\right) \exp(-\gamma(\delta^{-k}d(y_\alpha^k, x_0))^a). \\ &\lesssim \left(\frac{r}{r'}\right)^\eta \left(1 + \log_+ \frac{\tilde{r}}{r'}\right) \end{aligned}$$

*Conclusion.* Substituting (A.4), (A.5) and (A.6) into (A.3), we have shown that: for any  $B = B(x_0, r)$ , and  $\tilde{r} > r' > r$ , there exists a constant  $c = c(B, \tilde{r}, r')$  such that

$$\|1_B(b - c)\|_\infty \lesssim e^{-\gamma(\tilde{r}/r')^a} + \left(\frac{r}{r'}\right)^\eta \left(1 + \log \frac{\tilde{r}}{r'}\right).$$

If we now choose  $r' = tr$ ,  $\tilde{r} = tr' = t^2r$ , we get

$$\|1_B(b - c_{B,t})\|_\infty \lesssim e^{-\gamma t^a} + t^{-\eta}(1 + \log t) \xrightarrow[t \rightarrow \infty]{} 0.$$

This concludes the proof.  $\square$

## REFERENCES

- [1] AHLBERG, J. H., NILSON, E. N., AND WALSH, J. L. *The theory of splines and their applications*. Academic Press, New York, 1967.
- [2] AIMAR, H., CARENA, M., AND IAFFEI, B. Discrete approximation of spaces of homogeneous type. *J. Geom. Anal.* 19, 1 (2009), 1–18.
- [3] ALBEVERIO, S., EVDOKIMOV, S., AND SKOPINA, M.  $p$ -adic multiresolution analysis and wavelet frames. *J. Fourier Anal. Appl.* 16, 5 (2010), 693–714.
- [4] BATTLE, G. A block spin construction of ondelettes. I. Lemarié functions. *Comm. Math. Phys.* 110, 4 (1987), 601–615.
- [5] CHRIST, M. A  $T(b)$  theorem with remarks on analytic capacity and the Cauchy integral. *Colloq. Math.* 60/61, 2 (1990), 601–628.
- [6] COIFMAN, R. R., AND WEISS, G. *Analyse harmonique non-commutative sur certains espaces homogènes*. Lecture Notes in Mathematics, Vol. 242. Springer-Verlag, Berlin, 1971. Étude de certaines intégrales singulières.
- [7] DAHLKE, S., DAHMEN, W., WEINREICH, I., AND SCHMITT, E. Multiresolution analysis and wavelets on  $S^2$  and  $S^3$ . *Numer. Funct. Anal. Optimization* 16, 1-2 (1995), 19–41.
- [8] DAVID, G. *Wavelets and singular integrals on curves and surfaces*, vol. 1465 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1991.
- [9] DAVID, G., JOURNÉ, J.-L., AND SEMMES, S. Opérateurs de Calderón-Zygmund, fonctions para-accrétives et interpolation. *Rev. Mat. Iberoamericana* 1, 4 (1985), 1–56.
- [10] DEMKO, S. Inverses of band matrices and local convergence of spline projections. *SIAM J. Numer. Anal.* 14, 4 (1977), 616–619.
- [11] DENG, D., AND HAN, Y. *Harmonic analysis on spaces of homogeneous type*, vol. 1966 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2009. With a preface by Yves Meyer.
- [12] FEDERER, H. *Geometric measure theory*. Die Grundlehren der mathematischen Wissenschaften, Band 153. Springer-Verlag New York Inc., New York, 1969.
- [13] HAN, Y., AND LIN, C.-C. Algebra of Calderón-Zygmund operators on spaces of homogeneous type. *Taiwanese J. Math.* 7, 2 (2003), 309–328.



- [14] HAN, Y., MÜLLER, D., AND YANG, D. Littlewood-Paley characterizations for Hardy spaces on spaces of homogeneous type. *Math. Nachr.* 279, 13-14 (2006), 1505–1537.
- [15] HAN, Y., MÜLLER, D., AND YANG, D. A theory of Besov and Triebel-Lizorkin spaces on metric measure spaces modeled on Carnot-Carathéodory spaces. *Abstr. Appl. Anal.* (2008), Art. ID 893409, 250.
- [16] HAN, Y. S., AND SAWYER, E. T. Littlewood-Paley theory on spaces of homogeneous type and the classical function spaces. *Mem. Amer. Math. Soc.* 110, 530 (1994), vi+126.
- [17] HAN, Y. S., AND WEISS, G. Function spaces on spaces of homogeneous type. In *Essays on Fourier analysis in honor of Elias M. Stein (Princeton, NJ, 1991)*, vol. 42 of *Princeton Math. Ser.* Princeton Univ. Press, Princeton, NJ, 1995, pp. 211–224.
- [18] HEINONEN, J. *Lectures on analysis on metric spaces*. Universitext. Springer-Verlag, New York, 2001.
- [19] HYTÖNEN, T., AND KAIREMA, A. Systems of dyadic cubes in a doubling metric space. *Colloq. Math.* 126, 1 (2012), 1–33.
- [20] HYTÖNEN, T., LIU, S., YANG, D., AND YANG, D. Boundedness of Calderón-Zygmund operators on non-homogeneous metric measure spaces. *Canad. J. Math.*, to appear.
- [21] HYTÖNEN, T., AND MARTIKAINEN, H. Non-homogeneous  $Tb$  theorem and random dyadic cubes on metric measure spaces. *J. Geom. Anal.*, to appear.
- [22] KAIREMA, A. Two-weight norm inequalities for potential type and maximal operators in a metric space. *Preprint* (2011). arXiv:1105.2661.
- [23] LEMARIÉ, P. G. Ondelettes à localisation exponentielle. *J. Math. Pures Appl. (9)* 67, 3 (1988), 227–236.
- [24] LEMARIÉ, P. G. Base d'ondelettes sur les groupes de Lie stratifiés. (Basis of ondelettes on stratified Lie groups). *Bull. Soc. Math. Fr.* 117, 2 (1989), 211–232.
- [25] MACÍAS, R. A., AND SEGOVIA, C. Lipschitz functions on spaces of homogeneous type. *Adv. in Math.* 33, 3 (1979), 257–270.
- [26] MAGGIONI, M., AND MHASKAR, H. N. Diffusion polynomial frames on metric measure spaces. *Appl. Comput. Harmon. Anal.* 24, 3 (2008), 329–353.
- [27] MALLAT, S. G. Multiresolution approximations and wavelet orthonormal bases of  $L^2(\mathbf{R})$ . *Trans. Amer. Math. Soc.* 315, 1 (1989), 69–87.
- [28] MEYER, Y. Ondelettes et fonctions splines. In *Séminaire sur les équations aux dérivées partielles 1986–1987*. École Polytech., Palaiseau, 1987, pp. Exp. No. VI, 18.
- [29] MEYER, Y. Wavelets and operators. In *Analysis at Urbana, Vol. I (Urbana, IL, 1986–1987)*, vol. 137 of *London Math. Soc. Lecture Note Ser.* Cambridge Univ. Press, Cambridge, 1989, pp. 256–365.
- [30] MEYER, Y. *Ondelettes et opérateurs. II*. Actualités Mathématiques. [Current Mathematical Topics]. Hermann, Paris, 1990. Opérateurs de Calderón-Zygmund. [Calderón-Zygmund operators].
- [31] MITREA, D., MITREA, I., MITREA, M., AND MONNIAUX, S. A sharp metrization theorem for groupoids and applications to analysis on quasi-metric spaces. *Preprint* (2011).
- [32] NAZAROV, F., REZNIKOV, A., AND VOLBERG, A. The proof of  $A_2$  conjecture in a geometrically doubling metric space. *Preprint* (2011). arXiv:1106.1342.
- [33] NAZAROV, F., TREIL, S., AND VOLBERG, A. The  $Tb$ -theorem on non-homogeneous spaces. *Acta Math.* 190, 2 (2003), 151–239.
- [34] PALUSZYŃSKI, M., AND STEMPAK, K. On quasi-metric and metric spaces. *Proc. Amer. Math. Soc.* 137, 12 (2009), 4307–4312.
- [35] PEVNYI, A. B., AND ZHELUDEV, V. A. Construction of wavelet analysis in the space of discrete splines using Zak transform. *J. Fourier Anal. Appl.* 8, 1 (2002), 59–83.
- [36] SCHOENBERG, I. J. *Cardinal spline interpolation*. Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1973. Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics, No. 12.
- [37] STEIDL, G. Spline wavelets over  $\mathbf{R}, \mathbf{Z}, \mathbf{R}/N\mathbf{Z}$ , and  $\mathbf{Z}/N\mathbf{Z}$ . In *Wavelets: theory, algorithms, and applications (Taormina, 1993)*, vol. 5 of *Wavelet Anal. Appl.* Academic Press, San Diego, CA, 1994, pp. 155–177.
- [38] STRÖMBERG, J.-O. A modified Franklin system and higher-order spline systems on  $\mathbf{R}^n$  as unconditional bases for Hardy spaces. In *Conference on harmonic analysis in honor of Antoni Zygmund, Vol. I, II (Chicago, Ill., 1981)*, Wadsworth Math. Ser. Wadsworth, Belmont, CA, 1983, pp. 475–494.
- [39] SWELDENS, W. The lifting scheme: a construction of second generation wavelets. *SIAM J. Math. Anal.* 29, 2 (1998), 511–546.
- [40] TOLSA, X. BMO,  $H^1$ , and Calderón-Zygmund operators for non doubling measures. *Math. Ann.* 319, 1 (2001), 89–149.
- [41] VERDERA, J. The fall of the doubling condition in Calderón-Zygmund theory. In *Proceedings of the 6th International Conference on Harmonic Analysis and Partial Differential Equations (El Escorial, 2000)* (2002), no. Vol. Extra, pp. 275–292.
- [42] YANG, D., AND ZHOU, Y. Localized Hardy spaces  $H^1$  related to admissible functions on RD-spaces and applications to Schrödinger operators. *Trans. Amer. Math. Soc.* 363, 3 (2011), 1197–1239.
- [43] YANG, D., AND ZHOU, Y. New properties of Besov and Triebel-Lizorkin spaces on RD-spaces. *Manuscripta Math.* 134, 1-2 (2011), 59–90.

UNIV. PARIS-SUD, LABORATOIRE DE MATHÉMATIQUES, UMR 8628, F-91405 ORSAY; CNRS, F-91405 ORSAY  
*E-mail address:* `pascal.auscher@math.u-psud.fr`

DEPARTMENT OF MATHEMATICS AND STATISTICS, P.O. Box 68, FI-00014 UNIVERSITY OF HELSINKI, FINLAND  
*E-mail address:* `tuomas.hytonen@helsinki.fi`